Understanding Complex Systems

Springer : COMPLEXITY

Viorel Badescu

Modeling Thermodynamic Distance, Curvature and Fluctuations

A Geometric Approach

Springer

Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems—cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse "real-life" situations like the climate, the coherent emission of light from lasers, chemical reaction-diffusion systems, biological cellular networks, the dynamics of stock markets and of the Internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series "Understanding Complex Systems" focusing on the various applications of complexity, the "Springer Series in Synergetics", which is devoted to the quantitative theoretical and methodological foundations, and the "Springer Briefs in Complexity" which are concise and topical working reports, case studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Editorial and Programme Advisory Board

Understanding Complex Systems

Founding Editor: S. Kelso

Future scientific and technological developments in many fields will necessarily depend upon coming to grips with complex systems. Such systems are complex in both their composition – typically many different kinds of components interacting simultaneously and nonlinearly with each other and their environments on multiple levels – and in the rich diversity of behavior of which they are capable.

The Springer Series in Understanding Complex Systems series (UCS) promotes new strategies and paradigms for understanding and realizing applications of complex systems research in a wide variety of fields and endeavors. UCS is explicitly transdisciplinary. It has three main goals: First, to elaborate the concepts, methods and tools of complex systems at all levels of description and in all scientific fields, especially newly emerging areas within the life, social, behavioral, economic, neuro- and cognitive sciences (and derivatives thereof); second, to encourage novel applications of these ideas in various fields of engineering and computation such as robotics, nano-technology, and informatics; third, to provide a single forum within which commonalities and differences in the workings of complex systems may be discerned, hence leading to deeper insight and understanding.

UCS will publish monographs, lecture notes, and selected edited contributions aimed at communicating new findings to a large multidisciplinary audience.

More information about this series at http://www.springer.com/series/5394

Viorel Badescu

Modeling Thermodynamic Distance, Curvature and Fluctuations

A Geometric Approach

Springer

Viorel Badescu
Candida Oancea Institute
Polytechnic University of Bucharest
Bucharest
Romania

ISSN 1860-0832 ISSN 1860-0840 (electronic)
Understanding Complex Systems
ISBN 978-3-319-33788-3 ISBN 978-3-319-33789-0 (eBook)
DOI 10.1007/978-3-319-33789-0

Library of Congress Control Number: 2016939372

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

This textbook aims to briefly outline the main directions in which the geometrization of thermodynamics has been developed in the last decades. The textbook is accessible to the people trained in thermal sciences but not necessarily with solid formation in mathematics. For this, in the first part of the textbook a summary of the main mathematical concepts is made. In some sense, this makes the textbook self-consistent. The rest of the textbook consists of a collection of results previously obtained in this young branch of thermodynamics. The content is organized as follows.

The first part of the textbook, consisting of four chapters, presents the main mathematical tools. Thus, Chap. 1 presents the historical background of the geometrization of mechanics and thermodynamics. In Chap. 2 some basic concepts are briefly reminded, such as the set theory, the relationships theory, and the theory of simple algebraic structures. Then, the essential concepts used in the theory of linear spaces are introduced. The chapter ends by presenting some results concerning the coordinate transformations and the classification of physical quantities in relation with these transformations. Chapter 3 describes the main types of vectors and the standard method of vector geometrization. Then elementary results of vector calculus are presented. The chapter ends with a very brief introduction to the exterior differential calculus, accompanied by some specific useful results. Chapter 4 describes results of Riemann geometry. Two approaches are presented. The first one is the classic approach. The second approach is based on the theory of differential manifolds and tangent spaces. Both approaches allow defining the tensors of different orders, the Riemann metric and the covariant differentiation, among others. The parallel between the two approaches is very useful for a deeper understanding of concepts.

The second part of the textbook, consisting of five chapters, refers to the application of geometric methods in equilibrium thermodynamics. Chapter 5 summarizes some results of equilibrium thermodynamics. The approach based on potentials is presented, including the standard procedures using the energy representation and the entropy representation. Finally, the extreme principles and the mathematical conditions for thermodynamic stability are presented. Chapter 6 briefly shows some results of using tools of contact geometry in thermodynamics. Here only the first law of thermodynamics is geometrized. The chapter ends with a few examples of contact currents in thermodynamics. In Chap. 7 an approach based on statistical methods, which allows defining the notions of thermodynamic metric and thermodynamic distance, is presented. The second law of thermodynamics plays a key role in this context. The relationship between the thermodynamic distance and the entropy production is analyzed and links with the Gouy-Stodola theorem are highlighted. Horse-carrot type theorems are also introduced. The manner in which the thermodynamic curvature can be defined is exposed in Chap. 8. The chapter contains examples of calculation of thermodynamic curvature for simple systems. Chapter 9 presents a covariant theory of the thermodynamic fluctuations and analyzes the level of approximation introduced by the classical theory of fluctuations and its Gaussian approximation.

The textbook is a more extensive version of a section of the course of Advanced Thermodynamics presented for master students at the Faculty of Mechanical Engineering, Polytechnic University of Bucharest, starting from the 2003–2004 academic year. The textbook is presented with an ease of access for the readers with education in natural and technical sciences. Thus, most mathematical demonstrations of the theoretical results with higher degree of difficulty are omitted and references for the relevant literature are provided.

As usual, the preparation of such a work is the result of numerous interactions, discussions, consultations, and collaborations. It is a pleasure to remind here some of them. I received special support from colleagues in the European network CARNET (Carnot Network). This cooperation was institutionalized during the years 1994–1999 by two Copernicus projects on thermodynamic topics funded by the European Commission. In particular, I must thank Prof. Bjarne Andresen (University of Copenhagen), Prof. Ryszard Mrugala (University of Torun, Poland), and Dr. Lajos Diósi (Research Institute for Particle and Nuclear Physics, Budapest) whose publications were massively used in the present work. During the elaboration of the material I received technical support from Prof. Peter Salamon (University of San Diego). Also, discussions with Prof. Constantin Udriste (Polytechnic University of Bucharest) allowed a better understanding of the fundamentals of mathematics.

Viorel Badescu

Contents

Part I Mathematical Tools

Part II Applications of Geometric Methods in Thermodynamics

Part I
Mathematical Tools

Chapter 1
Introduction

Theoretical thermodynamics has been developed from the practical studies of thermal engines operation. Initially, it was based on the empirical usage of a combination of mechanical and thermal notions. The absence of a sound basis, consisting of well-defined and understood concepts, has often been noticed by personalities who made important contributions in the field, among which we quote Josiah Willard Gibbs, Hermann von Helmholtz, Pierre Duhem and Walther Nernst. They, and many others, have tried to introduce rigor in the theoretical approach by avoiding cyclical logical reasoning and contradictions. Constantin Caratheodory was the first who succeeded to build an axiomatic system for equilibrium thermodynamics (Caratheodory 1909). Thus, notions such as measurable temperature, heat and entropy were defined. Also, empirical assumptions were explained and simplified. From the very beginning, the approach of Caratheodory was analogous in spirit and practice with the axiomatic formulation of Euclidean geometry (Antoniou 2002). The structure of equilibrium thermodynamics, expressed in mathematical terms by Pfaff forms, turns out to be in some sense analogous to the structure of Hamiltonian mechanics and symplectic geometry (Rastal 1970; Peterson 1979).

The criteria proposed by Caratheodory were subsequently used to develop two main lines of research. First, the notion of thermodynamic variable has been clarified, making difference between extensive and intensive variables. Second, the concepts of classical thermodynamics have been extended for non-equilibrium situations.

Equilibrium thermodynamics was formulated based on conjugate pairs of independent variables, known as generalized coordinates (such as volume, area, length, electric charge) and generalized forces (mechanical forces, pressure, surface tension, electric voltage) (Redlich 1968). Generalized coordinates actually correspond to extensive variables because they depend on the size (extension) of the system, while the generalized forces correspond to the intensive variables because they are localized in space and time. This simple observation led to the formulation

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_1

of the fundamental thermodynamic concepts by using the measure theory (Gurtin et al. 1986). In this version, the extensive variables are positive or negative quantities on Euclidean spaces while the intensive variables are associated densities, defined mathematically as Radon-Nykodym derivatives. In case of extensive variables represented by absolute continuous quantities, the corresponding intensive variables are absolute integrable functions. In case of extensive variables represented by singular measures, located on surfaces, curves or fractals, the corresponding intensive variables are generalized functions (Antoniou and Suchanecki 1999). The formulation of classical thermodynamics by using of the measure theory has a number of advantages, among which we can mention the possibility of rigorous generalization to the case of continuous media, to the case of special and general relativity and to non-equilibrium situations, respectively (Antoniou 2002).

The usage of geometrical methods in thermodynamics was inspired by their previous applicability in the field of dynamical systems theory. The idea of approaching the solutions of the dynamic equations from a geometrical point of view is due to Henri Poincaré. However, Nikolay Mitrofanovich Krylov was the first who tried to formulate the statistical mechanics by using the Riemann geometry (see the review by Krylov (1979)). Krylov's ideas were developed by several groups of researchers (see Caiani et al. 1998; Casetti et al. 2000 and references therein). In general, conclusive results have been obtained only for constant negative curvature of the space of configurations. There are attempts to replace the Riemannian manifolds by Finsler manifolds, which have the advantage of allowing the geometrization of the speed dependent potentials (Dryuma 1994). Starting from the identification of the trajectories of a Hamiltonian dynamical system with geodesics in the configuration space equipped with Jacobi or Eisenhart metrics, one can develop a geometric theory of mechanics (Casetti et al. 2000). Interesting results have been obtained, which show, for example, that chaos can be induced not only by negative curvatures but also by positive curvatures of the configuration space, provided that these curvatures oscillate along the geodesics. In case of systems with very large number of particles and having large extension (what is commonly called "the thermodynamic limit") it is possible to describe the dynamical instability by using dynamic models that are independent of the dynamics of microparticles, which allows the analytical estimation of the largest Lyapunov coefficient as a function of the mean value and the fluctuations of the curvature of the configuration space. The main difficulty consists in the extremely complicated form of the geometry of the configuration space, in case of systems with many particles. Therefore, a number of more or less obvious simplifications are used in literature. Usually, these simplifications are a posteriori justified, by comparison with results obtained from computer simulations using statistical physics methods.

Applying geometric methods in thermodynamics was carried out mainly in the classical theory of equilibrium (see Ruppeiner 1991; Gross and Votyakov 2000). In this regard several procedures of geometrization have been proposed. Probably the most popular is the approach developed by Weinhold (1975). It relies on the fact that the differentials of the thermodynamic functions can be interpreted as vectors in a vector space. Then, one can propose a definition of the inner product on that vector space, in connection with the mathematical expression of the second law of thermodynamics, which ultimately leads to the positivity of the metrics attached to the vector space. The procedure initiated by Weinhold experienced many extensions, some of which will be mentioned throughout this book.

On the other hand, some geometrical aspects of the hypersurface of constant energy in the phase space were used in case of the microcanonical ensemble to define the temperature and the specific heat (Rugh 1997; Giardina and Livi 1998). Therefore, one can make distinction between the use of geometrical and topological concepts at the level of the macroscopic phase space (associated with phenomenological thermodynamics) and at the level of the microscopic phase space (associated with statistical thermodynamics), respectively.

When a system undergoes a phase change, fluctuations in the curvature of the phase/configurations space, as a function of temperature or energy, have a singular behavior in the transition point. This singularity can be described using a geometric model. In such a model the singularity of curvature fluctuations originates in the topology of the phase/configurations space. This is the argument leading to the introduction of the so-called topological assumption, which states that phase changes (at least, the continuous ones) are connected to a specific change of the topology of system's space of phase/configurations. This assumption allows the usage within the statistical thermodynamics of existing results in mathematics, such as those obtained in Morse theory. Therefore, one can make such a connection between mathematics (topology) and statistical thermodynamics (the theory of phase change). Existing results in the literature show that from the point of view of Morse theory the essential information is stored in the potential energy function. If the latter depends solely on coordinates, the usage of topological methods can be made by restriction from the phase space to the configurations space.

An important theorem shows the need of topological changes of the hypersurface of constant energy, for the emergence of a first-order or second-order phase change (Casetti et al. 2000). The demonstration is based on several assumptions concerning the diffeomorphicity of the surface and the uniform convergence of the Helmholtz free energy towards the thermodynamic limit for a very large number of particles. The fact that topological changes can occur regardless of the number of particles opens the possibility of describing the phase change in finite systems such as nuclear and atomic clusters, polymers and proteins, as well as nanoscopic and mesoscopic structures.

A geometric theory of thermodynamic fluctuations has already been proposed (for a review see, Ruppeiner 1995). The theory applies to classical, extensive, thermodynamics, in all cases where there are two independent coordinates (identified in that situation with two thermodynamic parameters). Then, a metric is defined on the manifold determined by the two coordinates, which becomes a Riemann manifold. Using known results of Riemann geometry is simplified in case of extensive thermodynamics with positive defined metrics, unlike in the theory of relativity (there, the associated Riemann manifold has four dimensions and its metric is allowed not to be positive definite). Arguments have been provided in support that the theory of fluctuations and, implicitly, the theory of thermodynamic stability, can be put into relation with the curvature (properly defined) of a Riemann manifold, which in that case has volume units. In case of the monatomic perfect gas it has also been showed that the curvature is null, due to the lack of interparticle interactions. The curvature tends towards infinity at critical points. The geometrization of thermodynamics in two dimensions was used to study the fluctuations in several cases of simple systems such as the Van der Waals gas, the paramagnetic systems, one-dimensional Ising models and quantum gases.

Important contributions to the geometrization of thermodynamics have brought Ryszard Mrugala and H. Janyszek (Mrugala 1978, 1984; Janyszek 1986, 1991; Janyszek and Mrugala 1989a, b) and the group of Peter Salamon (Salamon et al. 1985; Nulton and Salamon 1985). Also, by using geometric methods, Vsevolod Radcenco and Constantin Udriste have obtained the integral submanifold of the Gibbs-Pfaff equation, by defining ten types simple thermodynamic systems. The study was deepened by analyzing the minimum of the internal energy of thermodynamic systems, leading to the possibility of describing simple interactions of thermodynamic systems by using mathematical methods specific to the problems of extreme with non-holonomic constraints (Radcenco et al. 1991; Udriste 2000).

References

Antoniou, I.E.: Caratheodory and the foundations of thermodynamics and statistical physics. Found. Phys. **32**(4), 627–641 (2002)

Antoniou, I.E., Suchanecki, Z.: 1999. Densities of singular measures and generalized spectral decomposition. In: Antoniou, I., Lumer, G. (eds.) Generalized Functions, Operator Theory and Dynamical Systems. Research Notes in Mathematics, vol. 399, pp. 56–67. Chapman and Hall/CCR, Boca Raton (1999)

Caiani, L., Casetti, L., Pettini, M.: Topological origin of the phase transition in a mean-field model. J. Phys. A **31**, 3357 (1998)

Caratheodory, C.: Untersuchungen uber die Grundlagen der thermodynamik. Math. Ann. **67**, 355 (1909)

Casetti, L., Pettini, M., Cohen, E.G.D.: Geometric approach to Hamiltonian dynamics and statistical mechanics. Phys. Rep. **337**, 237–341 (2000)

Dryuma, V.S.: Geometrical properties of the multidimensional nonlinear differential equations and the Finsler metrics of phase spaces of dynamical systems. Teor. Mat. Fiz. **99**, 241 (1994)

Giardina, C., Livi, R.: Ergodic properties of microcanonical observables. J. Stat. Phys. **91**, 1027 (1998)

Gross, D.H.E., Votyakov, E.: Phase transitions in “small” systems. Eur. Phys. J. B **15**, 115 (2000). cond-mat/9911257

Gurtin, M., Williams, W., Ziemer, W.: Geometric measure theory and the axioms of continuum thermodynamics. Arch. Rat. Mech. Anal. **92**, 1–22 (1986)

Janyszek, H.: On the Riemannian metric structure in the classical statistical equilibrium thermodynamics. Rep. Math. Phys. **24**(1), 1–10 (1986)

Janyszek, H.: Riemannian geometry and stability of thermodynamical equilibrium systems. J. Phys. A: Math. Gen. **23**, 477–490 (1991)

Janyszek, H., Mrugala, R.: Geometrical structure of the state space in classical statisticsl and phenomenological thermodynamics. Rep. Math. Phys. **27**(2), 145–159 (1989a)

Janyszek, H., Mrugala, R.: Riemannian geometry and the thermodynamics of model magnetic systems. Phys. Rev. A **39**(12), 6515–6523 (1989b)

Krylov, N.S.: Works on the Foundations of Statistical Physics. Princeton Series in Physics. Princeton University Press, Princeton (1979)

Mrugala, R.: Geometrical formulation of equilibrium phenomenological. Thermodynamics. Rep. Math. Phys. **4**, 419 (1978)

Mrugala, R.: On the equivalence of two metrics in classical thermodynamics. Phys. A **125**, 631 (1984)

Nulton, J.D., Salamon, P.: Geometry of the ideal gas. Phys. Rev. A **31**, 2520 (1985)

Peterson, M.: Analogy between thermodynamics and mechanics. Am. J. Phys. **47**, 488–490 (1979)

Radcenco, V., Udriste, C., Udriste, D.: Thermodynamic systems and their interaction. Sci. Bull. I. P. B. Electr. Eng. **53**(3–4), 285–294 (1991)

Rastal, P.: Classical thermodynamics simplified. J. Math. Phys. **11**, 2955–2965 (1970)

Redlich, O.: Fundamental thermodynamics since Caratheodory. Rev. Mod. Phys. **40**, 556–563 (1968)

Rugh, H.H.: Dynamical approach to temperature. Phys. Rev. Lett. **78**, 772 (1997)

Ruppeiner, G.: Riemannian geometric theory of critical phenomena. Phys. Rev. A **44**, 3583 (1991)

Ruppeiner, G.: Riemannian geometry in thermodynamic fluctuation theory. Rev. Mod. Phys. **67**, 605–659 (1995)

Salamon, P., Nulton, J.D., Berry, R.S.: Length in statistical thermodynamics. J. Chem. Phys. **82**, 2433 (1985)

Udriste, C.: Geometric Dynamics. Kluwer Academic Publishers, Dordrecht (2000)

Weinhold, F.: Metric geometry of equilibrium thermodynamics. J. Chem. Phys. **63**, 2479 (1975)

Chapter 2
Algebraic Structures. Spaces. Reference Frames

Mathematical structures are usually defined by using sets of axioms. The definition of the axioms must meet the following three rules:

1. The sets of elements to which the axioms apply must not be empty.
2. The sets of elements to which the axioms apply must not be trivial; in other words, elements which do not fulfill the axioms must exist.
3. The axioms must be independent, i.e. none of the axioms should be obtained from the other axioms.

2.1 Sets

The definitions and the mathematical concepts are based on set theory. Moreover, the methods of mathematical thinking are combinations of arguments of mathematical logic and of set theory. To help the reader, some definitions and elementary results of the set theory are briefly presented (Gellert et al. 1980; Kaufmann and Precigout 1973). Georg Cantor (1845–1918) is the founder of the set theory. He gave the following definition: a set is obtained when several objects specified by human perception or by thought are included into a single entity; these objects are called the elements of the set. This definition, although imprecise and prone to induce some contradictions, has the advantage of an intuitive image. If an object x is an element of the set S, one writes $x \in S$. If S contains two distinct elements a and b, then S is called *unordered pair* and it is denoted $S = \{a, b\}$. A subset T of a set S is any set which contains only elements belonging to S. This is denoted $T \subseteq S$. It is said that the set T is included in the set S. The empty set is a set without elements. The set defined by a sentence $H(x)$ is denoted by $\{x \mid H(x)\}$ (It is read "the set of all x so that $H(x)$").

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_2

The sets whose elements are sets are called *families* (or systems, or classes) of sets. An important class is the set of all subsets of a given set S; this is called the power set of S (or the set of the parts of S) and is denoted $\mathbf{P}(S)$.

All the systems of axioms of the set theory have in common the following four principles.

- The *principle of extensionality,* which say that two sets S and T, having the same elements, are identical (it is written $S = T$).
- The *principle of construction* indicates different specific types of sentences that are used for defining the sets. Usually it requires that those sentences must contain only symbols of objects, logic symbols and the symbol $\in$.
- The *principle of the existence of infinite sets* must be understood as such. Although it is difficult to motivate it in connection with reality, without this principle an important part of mathematics (including the differential and the integral calculus) would lose its meaning.
- The fourth principle is usually called the *axiom of choice*: If S is a class of non-empty sets, then there is a set A that has precisely one element in common with each set of S.

For the construction of new sets, starting from given sets, operations with sets are used. The main operations are: union, intersection and difference, which are defined in Table 2.1.

Sets whose intersection is the empty set are called *disjoint sets*. If S is a subset of U, then U-S is called the *complement (or relative complement)* of S in U.

The main properties of operations with sets, which are often used in practice, are commutativity, associativity, distributivity and idempotency. For convenience, these well-known properties are reminded in Table 2.2.

Table 2.1 Operations with sets S and T

Operation name	Notation	Definition
Intersection	$S \cap T$	$\{x \mid x \in S \text{ and } x \in T\}$
Union	$S \cup T$	$\{x \mid x \in S \text{ or } x \in T\}$
Difference	$S \backslash T$	$\{x \mid x \in S \text{ and } x \notin T\}$

Table 2.2 Properties of operations with sets

Name	Explanation
Commutativity	$S \cap T = T \cap S \qquad S \cup T = T \cup S$
Associativity	$S \cap (T \cap R) = (S \cap T) \cap R \qquad S \cup (T \cup R) = (S \cup T) \cup R$
Distributivity	$S \cap (T \cup R) = (S \cap T) \cup (S \cap R)$ $S \cup (T \cap R) = (S \cup T) \cap (S \cup R)$
Idempotency	$S \cap S = S \qquad S \cup S = S$

If S and T are subsets of U and their complements in U are S' and T', respectively, then the following relations occur:

$$(S \cap T)' = S' \cup T' \qquad (S \cup T)' = S' \cap T' \tag{2.1}$$

These are the *laws of De Morgan*, often used in applications.

2.2 Relations

An *ordered pair* (a, b) is intuitively defined as a juxtaposition of two objects a and b so that a can be distinguished as the first element of the ordered pair and b as the second element. A rigorous definition will be given later.

A relation R on a set S is a set of ordered pairs of elements of S. If $(a, b) \in R$, it is said that R takes place for the ordered pair (a, b). Sometimes this is denoted aRb. For example, in the set S of all people alive at some moment, we can define the relationship "A is the parent of B".

The set $\{x \in S | (x, y) \in R$ for at least one y from $S\}$ is called the support of R. The set $\{y \in S | (x, y) \in R$ for at least one x from $S\}$ is called the set of values or the codomain (or the range, or the image) of R. These sets will be denoted by Sup R and Ran R, respectively. The set Dom $R =$ Sup R is called the domain of R. Of course, Dom $R \subseteq S$.

There are many relations that are commonly used in mathematics as well as in other areas, of which the exact sciences come on the first place. It is found, through a systematic analysis, that the relations have certain common properties, which are listed in Table 2.3.

Table 2.3 Attributes for a relationship R on a set S

Attribute	Definition
Reflexive	xRx takes place for all $x \in S$
Non-reflexive	Does not exist $x \in S$ so that xRx takes place
Symmetric	For every $x, y \in S$, from xRy it comes yRx
Asymmetric	Does not exist elements $x, y \in S$ with xRy and yRx
Antisymmetric	For every $x, y \in S$: if xRy and yRx, then $x = y$
Transitive	For every $x, y, z \in S$: if xRy and yRx, then xRz
Connex	For every $x, y \in S$: if $x \neq y$, then xRy or yRx
Left-unique (injective)	For every $x, y, z \in S$: if xRz and yRz, then $x = y$
Right-unique (univalent, right defined)	For every $x, y, z \in S$: if xRy and xRz, then $y = z$
Biunivocal (one-to-one)	If left-unique and right-unique

2.2.1 Equivalence Relations

An *equivalence relation* on a set S is a reflexive, symmetric and transitive relation which has S as support. For example, the relation of parallelism between two straight lines d si d', which is noted $d||d'$, is an equivalence relation.

An equivalence relation R on S induces a partition of S into classes, which are composed of those elements between which the equivalence relation is defined. A partition of a set S is a nonempty family **P** of non-vide subsets of S, called *partition classes*, with the following two properties:

- (i) two distinct classes are disjoint
- (ii) any element of S belongs to a class.

The next theorem is called the *principle of identification.*

Theorem 2.1 *If R is an equivalence relation on a set S, then there is a partition $\mathbf{P}$ of S so that the elements $(a, b) \in S$ are in the same class of $\mathbf{P}$, if, and only if aRb. Conversely, if $\mathbf{P}$ is a particular partition of S, then the relation $\{(a, b) |\, \text{there is a class } C \in P \text{ with } a, b \in C\}$ is an equivalence relation.*

2.2.2 Ordering

A relation R on the set S is called *partial order relation* if R is reflexive, transitive and antisymmetric. If, in addition, R is concave, the relation is called *total order relation* or *linear order relation*. For example, the relation $a \leq b$ is a partial order relation, actually a total order relation, on the set of real numbers.

An *ordered set* is defined as a pair (S, R), where R is a partial order relation on the set S. Often, for brevity, the ordered set is simply denoted S. On an ordered set it can be defined an upper bound and a maximal element. A lemma commonly used in mathematics (*Kuratowski-Zorn lemma*) states that, if a totally ordered set (S, R) has a upper bound on S, then S has a maximal element. One can prove that this lemma is equivalent to the axiom of choice.

The notion of ordered pair (a, b) can be rigorously defined as follows: $(a, b) \equiv \{\{a\}, \{a, b\}\}$. This definition specifies the difference between the positions of the two elements. Ordered pairs have the following fundamental property: $(a_1, a_2) = (b_1, b_2)$ if and only if $a_1 = b_1$ and $a_2 = b_2$.

Consider two sets S and T. The *Cartesian product* $S \times S$ of these sets (noted $S \times T$) is the set of all ordered pairs (a, b) with $a \in S$ and $b \in T$. The Cartesian product $S \times S$ is shortly noted S^2. The Cartesian product $S^2 \times S$ is noted S^3 and this system of notation can be generalized. The elements of S^n are called *n-tuple* of elements of S. For example the 3-tuple, $((a, b), c)$ also called *triplet* is simply denoted (a, b, c). Another example: the set of complex numbers can be considered as the Cartesian product $R \times R = R^2$ of the set of real numbers with itself.

A relation with *n arguments* or *n-ary* on S is defined as a subset of S^n. Relations with two arguments are called *binary relations.* Relations with n arguments are called *predicates*. For example, the relationship "point A lies between the points B and C" is a relation with three arguments for the points on a straight line.

2.3 Functions and Maps

A *function* on a set S with values in T is a right-unique relation with the support S and the set of values T. The term *map* (or *mapping*) is used to mean a function, sometimes with a specific property of particular importance. If the support is the entire set S, it is said that the *map* is of *S in T*. If set of values of a map is the entire set T, then it is said that the map is of *S on T*, or *surjective*. Examples of functions are the functions defined on the set of real numbers R with values in the same set (called *real functions*). The functions of n real variable are functions defined on R^n with values in R. The maps of the set of natural numbers N in itself are called *arithmetic functions*.

The functions on S with values in T are subsets of $S \times T$. In some branches of mathematics (e.g. complex analysis) functions are not defined as right-unique. The most commonly used functions are maps of a set S into another set T. The set of all maps of S in T is denoted T^S.

One can define functions or maps whose arguments are functions or maps. Such functions are called *operations* (e.g. maps of S^2 in S), *functionals* (functions defined on a set of functions with values in the set of real numbers), *operators* (functions defined on a set of functions with values in another set of functions), *functors* and *morphisms* (maps that preserve in some sense the algebraic structures).

If F is a function on S with values in T and if $(x, y) \in F$, then y is called the *image* of x by F or the *value* of F in x. This is denoted in various ways, such as, $y = x^F, y = xF, y = F(x)$ or $y = F_x$. The set $F^{-1}(y) = \{x \in S \mid F(x) = y\}$ is called the *inverse image* (or *preimage*) of y.

A function on S with values in T is called *injective*, *injection*, *invertible*, *one-to-one* or *biunivocal* if it is a left-unique relation. In this case, any element from the domain of the values has a unique image and the set $\{(y, x) \in T \times S | (x, y) \in F\}$ is a function on T with values in S, which is noted F^{-1} and is called the inverse function of F. If F is an injective function, then F^{-1} is a function if and only if F is surjective. Such a function is called bijective. The inverse of a bijective function is itself bijective.

2.4 Groups

An operation on a set S is an application which associates each ordered pair (a, b) of elements of S a third element c in the same set S. For example, ordinary addition and multiplication are operations on sets of integer, rational, real or complex numbers, respectively.

An operation (denoted by $\otimes$) on a set S is called *associative* if $(a \otimes b) \otimes c = a \otimes (b \otimes c)$ for any elements a, b and c of S. The operation is called commutative if $a \otimes b = b \otimes a$. An element e of the set S is called neutral element of the operation $\otimes$ if $a \otimes e = e \otimes a = a$ for any element a of S. The neutral element, if any, is unique. An element a' of S is called the *inverse* of the element a of S if $a' \otimes a = a \otimes a' = e$. Sometimes the notation a^{-1} is used to designate the inverse of the element a.

In case of operations similar with ordinary multiplication (called multiplicative-like operations), the neutral element is called *identity element.* In case of operations similar with ordinary addition (called additive-like operations), the neutral element is called *zero element* and the inverse element is called *additive inverse element* (or *opposite element*).

A group is a set G for which the following conditions are fulfilled:

1. An operation is defined on G;
2. That operation is associative;
3. The set G has a neutral element;
4. Any element of G has an inverse in G.

If the operation is commutative, the group is called *commutative group* (or *Abelian group*).

A group is *finite* or *infinite* as the set of its elements is finite or infinite. The number of elements is the *order of the group*.

A subset H of a group G is called subgroup if H is a group for the group operation defined on G. All groups with one element are called *trivial subgroups* and all subgroups of a group G different from G are called *proper subgroups*.

2.4.1 Homeomorphism

An mapping f of a group $(G, \otimes)$ in a group $(G', \circ)$ is called *homeomorphism* if the relation $f(a \otimes b) = f(a) \circ f(b)$ occurs for any elements $a, b \in G$. The left-side product is taken in G and the right-side product is taken in G'. The image of G by f is a subgroup of G'. If a surjective homeomorphism of G on G' exists, then G' is the homeomorphic image of G. An homeomorphism may apply distinct elements of G on the same element of G'. Homeomorphisms are not necessarily injective.

The definition of homeomorphism suggests that it preserves in a certain sense the structure of the original group. However, in general the image is "smaller" than the original group. The set of the elements of G applied on the neutral element of the image is a measure of the narrowing of G. These elements form a subgroup of G called the *kernel of the homeomorphism.*

2.4.2 Isomorphism

A bijective homeomorphism is called *isomorphism*. If f is an isomorphism of G on G', then its image is G'. If there is an isomorphism from G to G', then it is said that the groups G and G' are isomorphic.

The isomorphism is an *equivalence relation* between groups, so that the class of all groups is divided into isomorphism classes. Isomorphic groups have the same structure and the calculations follow the same laws, even if the elements are of different nature and operations are defined in different ways.

2.4.3 Automorphism

An automorphism is an isomorphism of the group G on itself. The composition of two automorphisms of a group G is also an automorphism of G. If f is an automorphism, then f^{-1} is also an automorphism. The automorphisms of G form a group for the operation of composition of functions. This group is called the *group of the automorphisms* on G.

2.5 Fields

A field is a set K of elements that meet the following conditions (axioms):

1. On K two operations are defined (they will be referred to as addition and multiplication).
2. The addition determines on K an Abelian group, with 0 being the neutral element.
3. The multiplication determines on the nonzero elements of K an Abelian group.
4. Multiplication is distributive in relation to addition; therefore, for any elements a, b, c of K the relation $a(b+c) = ab + ca$ is true.

Examples of fields (for which the two operations are the ordinary addition and multiplication, respectively) are the sets of rational numbers, real numbers and complex numbers. Intuitively, it can be said that a field is a set in which ordinary arithmetic operations can be performed. The fields are finite and infinite, according to their number of elements.

A subset P of a field K is called subfield if it satisfies the axioms of the field for the operations defined in K. Obviously, the sum and product of the elements of P should belong to P, as well as their inverse elements and their opposite (additive inverse) elements of P. The field K is called the *extension of the field P*.

Any field K can be considered as a *vector space*. For this, any subfield P can be considered a *set of scalars*. The addition is defined on the elements of the field as the operation of addition in the vector space, while the multiplication on the elements of the subfield is defined as the operation of multiplication by scalars. If the field K has a finite dimension n, the elements $\beta_1, \beta_2, \ldots, \beta_n$ can be found, so that any $\beta \in K$ is expressed by the unique form $\beta = c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$, where $c_1, c_2, \ldots, c_n$ are elements of P. The elements $\beta_1, \beta_2, \ldots, \beta_n$ form a *basis* of K on P.

Given two fields K_1 and K_2, a bijective mapping f from K_1 to K_2, with the properties that $f(a+b) = f(a) + f(b)$ and $f(a \cdot b) = f(a) \cdot f(b)$, for any elements $a, b \in K_1$, is called *isomorphism* of K_1 on K_2. In this case, K_1 and K_2 are called isomorphic. An isomorphism of the field K on itself is called *automorphism*.

2.6 Spaces

2.6.1 Linear Spaces

There are several ways to define and to present the main properties of the linear spaces. Here the guide is (Beju et al. 1976).

Two sets are given: a set V on which the operation + is defined and a set K on which two operations are defined, namely $\oplus$ and $\circ$. The three operations are called *internal composition laws*. It is assumed that that $(V, +)$ is an Abelian group and that $(K, \oplus, \circ)$ is a commutative field. An operation called "product" (denoted $*$) of the elements of the group V with elements of the field K, is defined as follows:

$$K \times V \ni (\alpha, \mathbf{x}) \to \alpha * \mathbf{x} \in V \tag{2.2}$$

The operation * is called *external composition law* on V, because it attaches the element $\alpha * \mathbf{x}$ of V to the couple consisting of the element α of K and the element $\mathbf{x}$ of V, respectively. Denote by $\mathbf{x}, \mathbf{y}$ two arbitrary elements of V, by α, β two arbitrary elements of K and by 1 the identity element of the filed K. It is assumed that the following four axioms are fulfilled:

$$\begin{aligned} \alpha * (\mathbf{x} + \mathbf{y}) &= \alpha * \mathbf{x} + \alpha * \mathbf{y} \\ (\alpha \oplus \beta) * \mathbf{x} &= \alpha * \mathbf{x} + \beta * \mathbf{x} \\ (\alpha \circ \beta) * \mathbf{x} &= \alpha * (\beta * \mathbf{x}) \\ 1 * \mathbf{x} &= \mathbf{x} \end{aligned} \tag{2.3a–d}$$

Note that Eq. (2.3a–d) regulates how the external composition law * operates together with the internal composition laws defined in V and K. Equation (2.3a) shows how the "product" * with elements of K behaves in relation to the addition of elements of V. Equation (2.3b) and (2.3c) show how the "product" * behaves in relation to the addition and multiplication of the elements of K. Equation (2.3d) defines the effect of the identity element of the field K in relation to the elements of the group V. This equation ensures that the set of the values of the external composition law equals the set V (with other words, the product with elements of the field K is surjective).

If the axioms (2.3a–d) are fulfilled, it is said that V is a linear space on the field K. It is denoted V/K. In practice, two particular cases are more important, namely when K is the field of real numbers and the field of complex numbers, respectively. It is said that V/K represents a real or a complex linear space, respectively.

For convenience, it is customary to denote by 0 and 1, the zero element and the identity element of the field K, respectively. By $\mathbf{0}$ is usually denoted the zero element of the group V.

2.6.1.1 Vectors and Scalars

If the space V/K is linear, the elements of the group V are called *vectors* and the elements of the field K are called *scalars*. For this reason, V/K is called *vector space*. Therefore, the operation * can be assimilated to the product between vectors and scalars. In this case, the zero element $\mathbf{0}$ is called *zero vector*.

2.6.1.2 Linear Subspace

Consider the linear space V/K and a subgroup U of V. Also, α and $\mathbf{x}$ are arbitrary elements of K and U, respectively. In case that $\alpha * \mathbf{x} \in U$, then U is a linear space on K, denoted U/K, and it is called *linear subspace* of V/K. The product with scalars of the elements of U is the product with scalars from V/K.

2.6.1.3 Linear Independence

Assume a linear space V/K and a number n of elements of V, i.e. $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in V$. The vectors $\mathbf{x}_i$ $(i = 1, \ldots, n)$ are called *linearly independent* if the relation

$$\alpha_1 \mathbf{x}_1 + \cdots \alpha_n \mathbf{x}_n = 0 \tag{2.4}$$

with $\alpha_i \in K(i = 1, \ldots, n)$, implies $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Otherwise, these vectors are called *linearly dependent*. An equivalent wording is: the system of vectors

$\{\mathbf{x}_i, i = 1, \ldots, n)\}$ is linearly independent (respectively, linearly dependent). In a linearly dependent system of vectors, one of the vectors (in general, not anyone) can be expressed as a linear combination of the other vectors.

2.6.1.4 Dimension of a Linear Space

The linear spaces can be divided into two types. Thus, if in a linear space the number of linearly independent vectors is infinite, the space is called of *infinite dimension*. These spaces are studied by the *functional analysis*.

A linear space V/K is called of *finite dimension* if there is a finite upper bound for the number of its linearly independent vectors. In other words, there is a natural number n, so that there are n linearly independent vectors, but, at the same time, any m vectors ($m > n$) are mandatory linearly dependent. The number n is called the *dimension of the linear space*.

2.6.1.5 Basis for a Linear Space

A maximal system of n linearly independent vectors constitutes a *basis* of V/K. In another formulation, if n linearly independent vectors can be found in a linear space, any $n+1$ vectors being linearly dependent, then the linear space has the dimension n (it is said that the space is n-dimensional), the n linearly independent vectors being its basis.

2.6.1.6 An Important Example of Linear Space

Assume a commutative field K. A set K^n is built, consisting of sets of n elements taken from K. Using the Cartesian product definition, the set K^n has the form:

$$K^n \equiv \underbrace{K \times K \times \cdots \times K}_{\text{n } times} \tag{2.5}$$

Two operations are defined: the "addition" of elements of K^n and the "multiplication" of elements of K with other elements of K^n:

$$\begin{aligned} (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) &\equiv (x_1 \oplus y_1, x_2 \oplus y_2, \ldots, x_n \oplus y_n) \\ \lambda * (x_1, x_2, \ldots, x_n) &\equiv (\lambda \circ x_1, \lambda \circ x_2, \ldots, \lambda \circ x_n) \end{aligned} \tag{2.6, 7}$$

The set K^n, for which the operations (2.6,7) are defined, has the structure of a linear space on the field K. This linear space, which is denoted K^n/K, is called n-dimensional Cartesian space, because its dimension is n. A basis of K^n/K is made up of the vectors

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, \ldots, 1) \tag{2.8}$$

Thus, any vector $\mathbf{x} \in K^n/K, x \equiv (x^1, x^2, \ldots, x^n)$ can be written as follows:

$$\mathbf{x} = (x^1, 0, \ldots, 0) + (0, x^2, \ldots, 0) + \cdots + (0, 0, \ldots, x^n) \tag{2.9}$$

which is equivalent with the more compact form:

$$\mathbf{x} = x^1 e_1 + x^2 e_2 + \cdots + x^n e_n \equiv x^i \mathbf{e}_i \tag{2.10}$$

The second equality in Eq. (2.10) represents the *Einstein's summation convention*, which is a compact way of writing some sums. From Eq. (2.10) it follows that any $n + 1$ vectors are linearly dependent.

2.6.1.7 Comments on the Axioms of Linear Spaces

Note that the set of axioms (2.3a–d) follows the rules presented in the beginning of this chapter. To show that the first rule is fulfilled, it must be proven that there are elements that meet the set of axioms. This can be easily proved by example (K^n/K is such an example of linear space). To show that the second rule, concerning the properties of the sets of axioms properly constructed, is checked, a counter-example will be given. Assume an additive Abelian group V and a field K. The following law of external composition with elements of V and K is defined:

$$\alpha * \mathbf{x} = \mathbf{0}, \quad (\forall \mathbf{x} \in V, \forall \alpha \in K) \tag{2.11}$$

Further assume two arbitrary elements $\mathbf{x}, \mathbf{y} \in V$ and two arbitrary elements $\alpha, \beta \in K$. Using the definition of the law of composition Eq. (2.11) it can be written:

$$\begin{aligned} \alpha * (\mathbf{x} + \mathbf{y}) &= \alpha * \mathbf{x} + \alpha * \mathbf{y} = \mathbf{0} \\ (\alpha \oplus \beta) * \mathbf{x} &= \alpha * \mathbf{x} + \beta * \mathbf{x} = \mathbf{0} \\ (\alpha \circ \beta) * \mathbf{x} &= \alpha * (\beta * \mathbf{x}) = \mathbf{0} \end{aligned} \tag{2.12a–c}$$

The three Eqs. (2.12a–c) lead to the following conclusions. The first conclusion is that the first three axioms (2.3a–c) of the definition of linear space are met. A second conclusion comes from the fact that the axiom Eq. (2.3d) is not met; therefore the second condition of the correct way of construction of the system of axioms is satisfied, because it was shown that there are elements that do not check the entire set of axioms. The third conclusion, which derives from the first two, is that axiom (2.3d) is independent of the other three axioms.

It may be shown that the axiom (2.3b) is independent, by using the following example, where V and K are an Abelian group and an arbitrary field, respectively. The product of the elements of V with scalars of K is defined as follows:

$$\alpha * \mathbf{x} = \mathbf{x}, \quad (\forall \mathbf{x} \in V, \forall \alpha \in K) \tag{2.13}$$

Assume two elements $\mathbf{x}, \mathbf{y} \in V$ and two elements $\alpha, \beta \in K$. Using the definition of the law of composition (2.12), it can be written:

$$\begin{aligned} \alpha * (\mathbf{x} + \mathbf{y}) &= \alpha * \mathbf{x} + \alpha * \mathbf{y} = \mathbf{x} + \mathbf{y} \\ (\alpha \circ \beta) * \mathbf{x} &= \alpha * (\beta * \mathbf{x}) = \mathbf{x} \\ 1 * \mathbf{x} &= \mathbf{x} \end{aligned} \tag{2.14}$$

It is found that the axioms (2.3a), (2.3c) and (2.3d) are checked. Instead, the axiom (2.3b) is not checked, because it is seen that

$$\begin{aligned} (\alpha \oplus \beta) * \mathbf{x} &= \mathbf{x}, \\ \alpha * \mathbf{x} + \beta * \mathbf{x} &= \mathbf{x} + \mathbf{x} \end{aligned} \tag{2.15}$$

Hence the axiom (2.3b) is independent of the other three axioms. The third rule of defining in a correct way a system of axioms asks the independence of the axioms (2.3a) and (2.3c). This can be done quite easily using procedures similar to those above.

2.6.1.8 Properties of Vector Spaces

Using the axioms of linear spaces, several important properties can be formulated. Demonstrations can be found in Beju et al. (1976, pp. 107–108).

1. For any $\mathbf{x} \in V$, it can be checked that $0 * \mathbf{x} = \mathbf{0}$.
2. For any $\alpha \in K$, it can be checked that $\alpha * \mathbf{0} = \mathbf{0}$.
3. In any linear space, $\alpha * \mathbf{x} = 0$ if and only if $\alpha = 0$ or $\mathbf{x} = \mathbf{0}$.
4. In any linear space, the axiom of the commutativity of the group (V, +) is a consequence of the axioms of the linear space.
5. Any vector system containing the zero vector is linearly dependent.
6. Any system of linearly independent vectors does not contain the zero vector.
7. A system consisting of a single vector is linearly independent if and only if $\mathbf{x} \neq \mathbf{0}$.

2.6.1.9 Coordinates in Linear Spaces

Assume an ordered basis $\mathbf{e}_1, \mathbf{e}_2,\ldots, \mathbf{e}_\mathrm{n}$ of an n-dimensional linear space V/K and an arbitrary element $\mathbf{x} \in V$. Note that the following $n+1$ vectors: $\mathbf{x}, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, are linearly dependent. It follows that some scalars $\alpha_0, \alpha_1, \ldots, \alpha_n \in K$ exist, which are not all null, so that

$$\alpha_0\mathbf{x} + \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n = 0 \tag{2.16}$$

It is mandatory that $\alpha_0 \neq 0$, because otherwise it would follow that the n vectors $\mathbf{e}_1, \mathbf{e}_2,\ldots, \mathbf{e}_\mathrm{n}$ are not linearly independent. Thus, the relation (2.16) can be written as:

$$\mathbf{x} = \lambda_1\mathbf{e}_1 + \lambda_2\mathbf{e}_2 + \cdots + \lambda_n\mathbf{e}_n \tag{2.17}$$

It is easily shown that the scalars $\lambda_i (i = 1, \ldots, n)$ are uniquely determined (Beju et al. 1976, p. 110). Therefore, a set of scalars $\lambda_i (i = 1, \ldots, n)$ is associated to any vector $\mathbf{x}$. This set of scalars is unique in a given ordered basis. They constitute the *coordinates* of $\mathbf{x}$ in that basis.

For this reason, an ordered basis in a linear space is called *coordinate system*.

2.6.1.10 Isomorphism of Linear Spaces

The above allow a one-to-one correspondence between the set of vectors and the set of rows (or columns) formed by using coordinates of vectors. By this correspondence, the operations of addition of vectors and the product of a vector by a scalar may be associated with operations in K^n/K by using the rows (or columns) of coordinates.

An important consequence of this observation is that, whatever the nature of the elements of a n-dimensional linear space V/K (these elements being functions, arrays, physical quantities, etc.), that space does not differ fundamentally (in terms of its operations) from the space K^n/K. This observation is more rigorously stated as follows:

Definition 2.1 Two vector spaces V/K and W/K are called isomorphic if there is a mapping (function) $f : V \to W$ which has the properties:

(a) the function f is bijective;
(b) for any elements $\mathbf{x}, \mathbf{y} \in V$ and $\lambda \in K$, the following relations are true

$$f(x+y) = f(x) + f(y), \qquad f(\lambda x) = \lambda f(x) \tag{2.18}$$

All linear spaces of dimension n, whatever their nature, are isomorphic among themselves and isomorphic with the space K^n/K, which can be imagined as a space of rows (or columns). In general, the group operations of different vector spaces differ between them. However, due to the isomorphism, they can be marked with the same sign (for example, +).

2.6.1.11 Scalar Product in Linear Spaces

Assume some elements $\mathbf{x}, \mathbf{y}, \mathbf{z}$ of the linear space V/K and the arbitrary element $\lambda \in K$. Also, it is denoted by $\overline{c}$ the complex number conjugate of c. The scalar product in the space V/K is defined as a mapping f that associates certain elements $\mathbf{x}, \mathbf{y}$ of the vector product $V \times V$ with an element, denoted $<\mathbf{x}, \mathbf{y}>$, of the field K, namely

$$f(x, y) \rightarrow \ <x, y> \tag{2.19}$$

and fulfills the following four axioms:

$$\begin{aligned}
&<\mathbf{x}, \mathbf{y}> \ = \ <\overline{\mathbf{y}, \mathbf{x}}> \\
&<\mathbf{x}+\mathbf{y}, \mathbf{z}> \ = \ <\mathbf{x}, \mathbf{z}> + <\mathbf{y}, \mathbf{z}> \\
&<\lambda \cdot \mathbf{x}, \mathbf{y}> \ = \lambda \cdot <\mathbf{x}, \mathbf{y}> \\
&\begin{cases} <\mathbf{x}, \mathbf{x}> \ > 0, & \text{if } \mathbf{x} \neq 0 \\ <\mathbf{x}, \mathbf{x}> \ = 0 & \text{if } \mathbf{x} = 0 \end{cases}
\end{aligned} \tag{2.20a–d}$$

Note that, in the particular case of real linear spaces ($K \equiv R$), the axiom (2.20a) becomes $<\mathbf{x}, \mathbf{y}> \ = \ <\mathbf{y}, \mathbf{x}>$, i.e. the scalar product is symmetric in its arguments. Using the definition of the scalar product it can be shown that the following general propositions take place:

$$\begin{aligned}
&<\mathbf{x}, \lambda \cdot \mathbf{y}> \ = \overline{\lambda} \cdot <\mathbf{x}, \mathbf{y}>, \quad <\mathbf{x}, \mathbf{y}+\mathbf{z}> \ = \ <\mathbf{x}, \mathbf{y}> + <\mathbf{x}, \mathbf{z}> \\
&<\mathbf{0}, \mathbf{x}> \ = 0, \quad (<\mathbf{x}, \mathbf{y}>)^2 \leq \ <\mathbf{x}, \mathbf{x}> \cdot <\mathbf{y}, \mathbf{y}>
\end{aligned} \tag{2.21a–d}$$

Here $\overline{\lambda}$ is the conjugate complex number of the number λ. In the case of real linear spaces ($K \equiv R$), $\overline{\lambda} = \lambda$. The proposition (2.21d) is the *Cauchy–Bunyakovsky–Schwarz inequality*.

2.6.1.12 Norm and Distance in Linear Spaces

The scalar product can be used to introduce the notions of norm and distance on a linear space. The norm of an element $\mathbf{x}$ of V/K, usually denoted by $\|\mathbf{x}\|$, is a function from the set V/K in the set of real numbers, defined by

$$\|\mathbf{x}\| \equiv (<\mathbf{x},\mathbf{x}>)^{1/2} \tag{2.22}$$

The norm on a linear space fulfills the standard axioms of the norm-like functions

$$\begin{aligned} &\begin{cases} \|\mathbf{x}\| \geq 0 & \text{if} \quad \mathbf{x} \neq \mathbf{0} \\ \|\mathbf{x}\| = 0 & \text{if} \quad \mathbf{x} = \mathbf{0} \end{cases} \\ &\lambda \cdot \|\mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\| \\ &\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \end{aligned} \tag{2.23a–c}$$

Relation (2.23c) is the *triangle inequality.*

A vector $\mathbf{u} \in E$ is called *unit vector* or *normalized vector*, or *versor*, if $\|\mathbf{u}\| = 1$. Any vector $\mathbf{x} \neq \mathbf{0}$ can be normalized by dividing it by $\|\mathbf{x}\|$.

The distance between the elements $\mathbf{x}$ and $\mathbf{y}$ of V/K, which is usually denoted by $d(\mathbf{x},\mathbf{y})$, is a function from the set $V \times V$ in the set of real numbers, defined by:

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \tag{2.24}$$

The distance between two elements in a linear space fulfills all standard axioms of the distance-like functions, i.e.:

$$\begin{aligned} &\begin{cases} d(\mathbf{x},\mathbf{y}) \geq 0 & \text{if} \quad \mathbf{x} \neq \mathbf{y} \\ d(\mathbf{x},\mathbf{y}) = 0 & \text{if} \quad \mathbf{x} = \mathbf{y} \end{cases} \\ &d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x}), \qquad d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}) \geq d(\mathbf{x},\mathbf{z}) \end{aligned} \tag{2.25a–c}$$

It is seen from property (2.25b) that the distance is a symmetric function in its arguments.

2.6.2 Unitary and Euclidean Spaces

The concepts of scalar product, norm and distance, previously defined for linear spaces, have properties similar with the analogous notions of the Euclidean geometry. This observation allows an extension of the formal analogy, by introducing new definitions.

The complex linear spaces in which a scalar product has been defined are called *unitary spaces* while the real linear spaces in which a scalar product has been defined are called *Euclidean spaces.* The dimension of the unitary (Euclidean) space is given by the dimension of the linear space V/K. The Euclidean spaces are usually denoted by E, while the scalar product in Euclidean spaces is usually written in the more compact form $\mathbf{x} \cdot \mathbf{y}$, commonly used in geometry. Several notions of geometric inspiration will be introduced and discussed further.

2.6.2.1 Orthogonal Vectors

In the Euclidean space E, the scalar product can be used to introduce the notion of angle between two vectors. Assume $\mathbf{x}$ and $\mathbf{y}$ are two non-zero vectors. The angle $\theta(\mathbf{x},\mathbf{y})$ between these vectors is defined by the relation:

$$\cos\theta(\mathbf{x},\mathbf{y}) = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|}, \qquad (0 \le \theta \le \pi) \tag{2.26}$$

By using the Cauchy–Bunyakovsky–Schwarz inequality and the properties of the cosine function it can be shown that the angle $\theta(x,y)$ is well defined, i.e. it has values in the range [0,1].

Two vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if their scalar product is zero, i.e.:

$$\mathbf{x}\cdot\mathbf{y} = 0 \tag{2.27}$$

By using relations (2.26) and (2.27), it can be easily deduced that the angle between two orthogonal vectors equals $\pi/2$.

A set of nonzero vectors $\mathbf{x}_i (i = 1,\ldots,n)$ is called *orthogonal system* if $\mathbf{x}_i \cdot \mathbf{x}_j \neq 0$ only for $i = j, (i, j = 1,\ldots,n)$. The systems of orthogonal vectors have several important properties that result from the following two theorems (for demonstration, see Beju et al. 1976, p. 117):

Theorem 2.2 *An orthogonal system of vectors* $\{\mathbf{x}_i, i = 1,\ldots,n\}$ *is linearly independent.*

Theorem 2.3 *In any Euclidean space of finite dimension there are orthogonal bases.*

2.6.2.2 Orthogonalization Process

Assume a basis $\{\mathbf{e}_i, i = 1,\ldots,n\}$ of a n-dimensional Euclidean space E. Obviously, $\mathbf{e}_i \neq 0\,(i = 1,\ldots,n)$. Starting from this basis, an orthogonal basis (denoted $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n$) can be obtained by using the following procedure, called *orthogonalization* or *orthogonalization process.*

First, denote $\mathbf{f}_1 = \mathbf{e}_1$. Then, define $\mathbf{f}_2 = \mathbf{e}_2 + \alpha\cdot\mathbf{f}_1$ and determine the scalar α, by requiring that $\mathbf{f}_1$ and $\mathbf{f}_2$ are orthogonal (i.e. $\mathbf{f}_2\cdot\mathbf{f}_1 = 0$). It is found that $\alpha = -\mathbf{e}_2\cdot\mathbf{f}_1/\|\mathbf{f}_1\|^2$. Next, the mathematical induction is used. It starts from the premise that a set of k non-zero vectors, orthogonal two by two, denoted $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n$, has been built by using the relation

$$\mathbf{f}_k = \mathbf{e}_k + \alpha_1 \cdot \mathbf{f}_1 + \cdots + \alpha_{k-1} \cdot \mathbf{f}_{k-1} \qquad (k = 1, \ldots, m-1) \tag{2.28}$$

In this case, any vector $\mathbf{f}_m$ will be given by a relation similar with equality (2.28), i.e.

$$\mathbf{f}_m = \mathbf{e}_m + \alpha_1 \cdot \mathbf{f}_1 + \cdots + \alpha_{m-1} \cdot \mathbf{f}_{m-1} \tag{2.29}$$

The scalar coefficients $\alpha_1, \ldots, \alpha_{m-1}$ will be determined by using the orthogonality conditions $\mathbf{f}_m \cdot \mathbf{f}_k = 0 (k = 1, \ldots, m-1)$. In the generic case, it is found that $\mathbf{e}_m \cdot \mathbf{f}_k + \alpha_k \cdot \mathbf{f}_k \cdot \mathbf{f}_k = 0$, from which the coefficient α_k is determined.

A orthogonal system of normalized vectors is called *orthonormal system*. From Theorem 2.3 it follows that any n-dimensional Euclidean space has *orthonormal bases*. Note that in an orthonormal basis, the scalar product has the simple form:

$$\mathbf{x} \cdot \mathbf{y} = \sum_i x_i y_i \equiv x_i y_i \tag{2.30}$$

where $x_i, y_i (i = 1, \ldots, n)$ are the components of the vectors $\mathbf{x}$ and $\mathbf{y}$, respectively, on the orthonormal basis $\{\mathbf{e}_i\}$.

2.6.3 Affine Spaces

Geometry was the first branch of mathematics where sets of vectors have been attached to sets of points. With the generalization of the notion of vector for linear spaces, this observation is valid in the case of several disciplines of physics, such as mechanics, electrodynamics and, as will see in the next chapters, thermodynamics. For example, the force vector can be attached to a material point and a velocity vector field can be attached to a continuous domain of points of the physical space (i.e. a three-dimensional body).

The notion of affine space allows processing within the same mathematical structure of two distinct categories of elements, some of which are described as "points" and the other are called "vectors." The two categories of elements include geometric points and geometric vectors as particular cases.

The set of points will be denoted by M. The linear space of the vectors will be denoted V. Each *ordered pair* of points of M (for example, P, Q) is associated with a vector of V (for example $\mathbf{x}$) by defining an *association law*. In this case, the first point, P, is called the origin, or initial point, of the vector $\mathbf{PQ} \equiv \mathbf{x}$, and the second point, Q, is called its terminal point. With this notation, it is obvious that $\mathbf{PQ} \in V$.

For the set M, associated with the linear space V, to constitute an affine space, the following axioms should be fulfilled:

(a) For any point $\mathrm{P} \in M$ and any vector $\mathbf{x} \in V$, there is a point, and only one, $\mathrm{Q} \in M$, so that $\mathbf{PQ} = \mathbf{x}$.

(b) If $\mathbf{PQ} = \mathbf{x}$ and $\mathbf{QR} = \mathbf{y}$, then $\mathbf{PR} = \mathbf{x} + \mathbf{y}$.
Note that the ordered pair of overlaid points $(\mathrm{P,P}) \in M$ is associated with the null vector of V, i.e. $\mathbf{PP} = \mathbf{0}$. Also, if $\mathbf{PQ} = \mathbf{x}$, then it can be deduced that $\mathbf{QP} = -\mathbf{x}$ (Beju et al. 1976, p. 113).

If the linear space V is real or complex, finite or infinite dimensional, then it is said that the affine space is, respectively, real or complex, finite or infinite dimensional. The dimension of the affine space M is equal to the dimension of the linear space V.

2.6.3.1 Coordinates in Affine Spaces

Assume the n-dimensional affine space M. Assume a point O in M and a basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in the linear space V. The assembly consisting of the point O and the basis constitutes a system of affine coordinates in M. The point O is the origin of the coordinate system and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the vectors of the basis.

Assume a point P of M. The ordered pair $(\mathrm{O,P})$ is associated with the vector $\mathbf{OP}$, which is called the *position vector* of P in relation with the origin O. The coordinates of $\mathbf{OP}$ in the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are denoted $(x_1, x_2, \ldots, x_n)$. In this case, the following relation can be written:

$$\mathbf{OP} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \equiv x_i\mathbf{e}_i \tag{2.31}$$

The coefficients $(x_1, x_2, \ldots, x_n)$ represent the *affine coordinates* of the point P. Note that these coordinates depend on the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ of the linear space V and the origin O (element of M).

Assume the point $\mathrm{Q} \in M$, of affine coordinates $y_1, y_2, \ldots, y_n$. The affine coordinates of the vector $\mathbf{PQ}$ are given by:

$$\mathbf{PQ} = \mathbf{PO} + \mathbf{OQ} = \mathbf{OQ} - \mathbf{OP} = (y_1 - x_1)\mathbf{e}_1 + \cdots + (y_n - x_n)\mathbf{e}_n \tag{2.32}$$

Relation (2.32) was obtained by considering the axiom (b) of the affine spaces and the fact that $\mathbf{PO} = -\mathbf{OP}$. The conclusion is that the affine coordinates of the vector $\mathbf{PQ}$ can be written using only the affine coordinates of the points P and Q.

2.6.3.2 Distance in Affine Spaces

Consider the case of an affine space which is defined by using a linear space on which a scalar product has been introduced. In this case, the distance between two points A and B of the affine space can be defined by the relationship

$$d(\mathrm{A},\mathrm{B}) = (\mathbf{AB} \cdot \mathbf{AB})^{1/2} \tag{2.33}$$

By choosing an affine coordinate system consisting of a point O and an orthonormal basis, one obtains an image of the affine space similar to that of the space in Euclidean geometry.

2.6.3.3 Connection Between Affine Spaces and Linear Spaces

Between affine spaces and linear spaces there is a structural link. Thus, any affine space *M* can be regarded as a linear space *V*. To prove this, it takes an arbitrary point $\mathrm{O} \in M$. Next, a position vector **OP** is attached to any point $\mathrm{P} \in M$. Then, considering the axioms (a) and (b), it is concluded that the set of position vectors coincides with *V*. Conversely, any linear space *V* can be regarded as an affine space *M*. To prove this, the elements of *V* are regarded as points of *M* and to the ordered pair of points (a, b) it is attached the vector $\mathbf{b} - \mathbf{a} \in V$. This interpretation is encountered when making reference to the linear space of position vectors, defining, with its help, the vectors-oriented segments connecting the ends of two vectors of position.

To remove ambiguities, it is important to specify which of the analyzed elements have the role of points and vectors, respectively.

2.7 Equivalence Classes for Reference Frame Transformation

Practice shows that the same physical quantity may be perceived and described differently by different observers. The objective (scientific) description of the quantity must be done using the intrinsic properties of that quantity, which do not change when passing from the description of one observer to another description, of a different observer. More generally, the objective character of scientific knowledge can only be ensured by consistent usage of the invariant aspects of the phenomena, which do not depend on how the phenomena are perceived by particular observers.

In this section it is exemplified how the issues which are invariant in respect to the observer can be described by using mathematical methods. The simplest example refers to a fundamental concept in physics, that of distance. At the end of the section some useful generalizations are presented.

2.7.1 Intrinsic Distance

It is accepted that the mathematical structure of the physical space is that of the Euclidean space, denoted E^3. Further, by *observer* it is understood a right-handed orthonormal frame (for definition, see Sect. 3.1.1.1) equipped with a procedure for measuring distances and time.

Assume two observers, denoted $ox_1x_2x_3$ and $OX_1X_2X_3$, respectively. Assume two distinct points, R and S, of coordinates $\mathbf{y} \equiv (y_1, y_2, y_3)$ and $\mathbf{z} \equiv (z_1, z_2, z_3)$, respectively, in respect with the first observer, and of coordinates $\mathbf{Y} \equiv (Y_1, Y_2, Y_3)$ and $\mathbf{Z} \equiv (Z_1, Z_2, Z_3)$, respectively, in respect with the second observer. The distances between the points R and S, determined by the two observers, are given by, respectively:

$$\begin{aligned} d^2(\mathbf{y},\mathbf{z}) &= (z_1 - y_1)^2 + (z_2 - y_2)^2 + (z_3 - y_3)^2 \\ D^2(\mathbf{Y},\mathbf{Z}) &= (Z_1 - Y_1)^2 + (Z_2 - Y_2)^2 + (Z_3 - Y_3)^2 \end{aligned} \qquad (2.34\text{a, b})$$

It is clear that, generally, the functions $d(\mathbf{y},\mathbf{z})$ and $D(\mathbf{Y},\mathbf{Z})$ are different. Imposing further restrictions makes possible that these functions always lead to the same numerical value, once the points R and S are given. That numerical function, which does not depend on the procedures adopted by observers for measuring the distance, but only on the position of the two points, is called *intrinsic distance*.

Of interest are the conditions that must be met in order to obtain intrinsic distances from the relationship of type (2.34a, b). Denote by P a certain point, of coordinates $\mathbf{x} \equiv (x_1, x_2, x_3)$ in respect to the first observer, and of coordinates $\mathbf{X} \equiv (X_1, X_2, X_3)$ in respect to the second observer. Denote by f the function that makes the connection between coordinates $\{x_i\}$ and $\{X_i\}$ of the point P, as determined by the two observers. Obviously, this function performs a biunivocal correspondence between E^3 and E^3. In vector notation, it be written:

$$\mathbf{X} = f(\mathbf{x}) \qquad (2.35)$$

In case that the coordinates of the position vectors $\mathbf{x}$ and $\mathbf{X}$ of point P are used, the relationship (2.35) becomes

$$X_K = f_K(x_1, x_2, x_3) = f_K(x_i) \qquad (K, i = 1, 2, 3) \qquad (2.36)$$

The condition of existence of an intrinsic distance in space E^3 is equivalent to the equality

$$d(\mathbf{y},\mathbf{z}) = D(\mathbf{Y},\mathbf{Z}) \qquad \forall R, S \in E^3 \qquad (2.37)$$

The constraint (2.37) strongly restricts the set of functions f that fulfill the relationship (2.36). Now, it is considered that, for the first observer, the three points P, R and S are collinear. This is equivalent with writing

$$d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}) = d(\mathbf{x},\mathbf{z}) \tag{2.38}$$

Restriction (2.37) leads to the following equality

$$D(\mathbf{X},\mathbf{Y}) + D(\mathbf{Y},\mathbf{Z}) = D(\mathbf{X},\mathbf{Z}) \tag{2.39}$$

which is equivalent to saying that the points P, R and S are collinear from the view-point of the second observer. Therefore, the bijective function f transforms collinear points into collinear points, or, in other words, it transforms straight lines into straight lines. It follows that the function f must be a linear function, according to a theorem of affine geometry (see Mihaileanu 1971). It can be concluded that the relationships between the components of the position vectors $\mathbf{x}$ and $\mathbf{X}$ of the point P in the frames of the two observers must have the following form:

$$X_K = Q_{Kk}x_k + B_K \qquad (K = 1,2,3) \tag{2.40}$$

where Q_{Kk} and B_K $(K, k = 1,2,3)$ are constants. Denote by $\mathbf{Q}$ the matrix of the components Q_{Kk} and by $\mathbf{B}$ the column matrix of the components B_K. The position vectors $\mathbf{x}$ and $\mathbf{X}$ will be considered of column matrix type. In these conditions, Eq. (2.40) can be written under the following matrix form, which emphasizes the character of linear transformations

$$\mathbf{X} = \mathbf{Q}\cdot\mathbf{x} + \mathbf{B} \tag{2.41}$$

Classical mechanics, for example, is based on such relationships for changing the coordinates of a point from a reference frame into another. The linearity of the relationship (2.41) is a consequence of the premise that an intrinsic distance exists in the physical space modeled by E^3.

2.7.2 Orthogonal Transformations

It was shown that the distance between two points in the physical space E^3 is invariant for the coordinate transformations (2.41). Therefore, these so-called *orthogonal transformations* are important, they being able to underpin an objective description of the physical phenomena.

The orthogonal transformations are of two types, as seen below. Assume two points $\mathbf{y},\mathbf{z} \in E^3$. By using relations (2.34) and (2.37) it is obtained:

$$(Z_K - Y_K)(Z_K - Y_K) = (z_k - y_k)(z_k - y_k) \tag{2.42}$$

Then, using the relation (2.40), it can be written:

$$Z_K - Y_K = Q_{Kk}(z_k - y_k) \qquad (K = 1, 2, 3) \tag{2.43}$$

Therefore, the Eq. (2.42) becomes

$$Q_{Kk}(z_k - y_k)Q_{Kj}(z_j - y_j) = (z_k - y_k)(z_j - y_j) \tag{2.44}$$

By the reversal of the summing order, from (2.44) it is obtained:

$$(Q_{Kk}Q_{Kj} - \delta_{kj})(z_k - y_k)(z_j - y_j) = 0 \tag{2.45}$$

In (2.45) the symbol δ_{ij} has been used, which is known as *Kronecker's symbol*, defined as:

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \tag{2.46}$$

The equality (2.45) should take place regardless of the elements $\mathbf{y}, \mathbf{z} \in E^3$. Therefore, the following relation should be true:

$$Q_{Kk}Q_{Kj} = \delta_{kj} \qquad (k, j = 1, 2, 3) \tag{2.47}$$

Relation (2.47) can be more compactly rewritten in the following matrix form:

$$\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{E} \tag{2.48}$$

In the equality (2.48) the common notation has been used: $\mathbf{Q}^T$ is the transposed matrix of the matrix $\mathbf{Q}$, and $\mathbf{E}$ is the unit matrix. All these matrices are of order three. A consequence of the relationship (2.48) is that

$$\det(\mathbf{Q}^T \cdot \mathbf{Q}) = 1 \tag{2.49}$$

or, in other words, that

$$(\det \mathbf{Q})^2 = 1 \tag{2.50}$$

The previous relationships allow some comments about important properties of the matrix $\mathbf{Q}$. First, physical quantities are generally characterized by their physical dimension (e.g., dimension of length (L), mass (M), time (T)). In some cases, the physical dimension may be a more complex expression, of the form $L^a M^b T^c$, where a, b, c are real numbers. However, from the relation (2.41) it is found that the matrix $\mathbf{Q}$ has dimensionless components from physical point of view. Second, the matrix $\mathbf{Q}$ is non-singular. Then, from relations (2.48) and (2.50) it is found that the matrix $\mathbf{Q}$ is invertible. The inverse of the matrix $\mathbf{Q}$ is:

$$\mathbf{Q}^{-1} = \mathbf{Q}^T \tag{2.51}$$

Using again the eq. (2.48) it is deducted that:

$$\mathbf{Q} \cdot \mathbf{Q}^T = \mathbf{E} \tag{2.52}$$

or, in unfolded writing:

$$Q_{Kk}Q_{Lk} = \delta_{KL} \qquad (K, L = 1, 2, 3) \tag{2.53}$$

The matrices that check the equalities (2.51) are called *orthogonal matrices*. This explains why the linear relationships (2.41), where the matrix Q appears, are called orthogonal transformations. It is also observed that if in (2.41) an orthogonal matrix Q is included, the relationship (2.37) is checked unconditionally. This is justified by the next calculations, if the relationship (2.47) is also taken into account:

$$\begin{aligned} D^2(\mathbf{Y}, \mathbf{Z}) &= (Z_K - Y_K)(Z_K - Y_K) = Q_{Kk}(z_k - y_k)Q_{Kj}(z_j - y_j) \\ &= Q_{Kk}Q_{Kj}(z_k - y_k)(z_j - y_j) = \delta_{jk}(z_k - y_k)(z_j - y_j) \\ &= (z_k - y_k)(z_k - y_k) = d^2(\mathbf{y}, \mathbf{z}) \end{aligned} \tag{2.54}$$

Relation (2.50) can be rewritten in the following form, which is useful for the classification of the orthogonal matrices and transformations:

$$\det \mathbf{Q} = \pm 1 \tag{2.55}$$

An orthogonal transformation is said to be proper transformation, if $\det \mathbf{Q} = 1$. In the opposite case ($\det \mathbf{Q} = -1$), the transformation is called improper. Two reference frames (or two observers) are said to belong to the same class, if the orthogonal transformation that turns one into the other is a proper orthogonal transformation. Otherwise, it says that those reference frames (or observers) belong to different classes. From Eq. (2.55) one sees that there are only two classes of reference frames. The reference frames in the same class turns one into another by using a proper orthogonal transformation. The reference frames in different classes turns one into another by using an improper orthogonal transformation.

2.7.3 Classes of Physical Quantities

The importance of the transformations (2.40) or (2.41) between two reference frames $ox_1x_2x_3$ and $OX_1X_2X_3$, respectively, is that they lead to an intrinsic definition of the length. This has consequences on the shape and size of bodies.

The procedure developed previously for the intrinsic definition of the length can be extended to other types of physical quantities. Several types of physical quantities will be defined and classified in the following, according with their behavior when the transformation (2.40) occurs.

2.7.3.1 Scalars. Pseudoscalars. Vectors. Pseudovectors

In the following, only quantities with physical dimension will be considered. Therefore, the existence of this property will not be reminded.

A quantity characterized in any reference frame by a single real number invariant to the transformation (2.41) is called *scalar*. It is called *pseudoscalar*, a quantity characterized in any reference frame by a real number, which, when the transformation (2.41) occurs, is changed according to the rule:

$$\Omega = (\det Q)\omega \tag{2.56}$$

where ω and Ω is the value of the pseudoscalar in the reference frames $ox_1x_2x_3$ and $OX_1X_2X_3$, respectively.

A quantity characterized in any reference frame by a triplet of real numbers is called a *vector* if, when the transformation (2.41) occurs, the following relation is fulfilled between the triplets (v_1, v_2, v_3) and (V_1, V_2, V_3) which characterize the quantity in the two reference frames:

$$V_K = Q_{Kk}v_k \qquad (K = 1, 2, 3) \tag{2.57}$$

The three numbers represent the *vector components* in that reference frame. The vector thus defined is called a *polar vector*. A quantity characterized in any reference frame by a triplet of real numbers is called *pseudovector* if, when the transformation (2.41) occurs, the following relation is fulfilled between the triplets (v_1, v_2, v_3) and (V_1, V_2, V_3) which characterize the quantity in the two reference frames:

$$V_K = (\det Q)Q_{Kk}v_k \qquad (K = 1, 2, 3) \tag{2.58}$$

The three numbers represent the components of the pseudovector in that reference frame. The pseudovector is also called *axial vector*.

2.7.3.2 The Importance of Physical Dimension

The quantities previously defined (scalar, pseudoscalar, vector, pseudovector) need not be associated with a physical dimension, as long as they are maintained at an abstract level. Sometimes, using these parameters in practice requires specification of physical dimension. In this way it is avoided the composition of quantities that have obviously different physical significance (such as velocities and accelerations) although they are of the same type (i.e. vectors, in this particular case). As already mentioned, the quantities Q_{Kk} are dimensionless. By assuming the homogeneity of the transformations (2.56) and (2.58), it is seen that the physical dimension of the quantities is kept when the reference frame is changed.

2.7.3.3 Examples of Vectors

Next, some examples of mechanical vectors are shown. The theory can easily be generalized and used in other areas of physics.

Assume a reference frame in the three-dimensional space and the points P^0 (of coordinates $\left(x_1^0, x_2^0, x_3^0\right)$ in that reference frame) and P (of coordinates (x_1, x_2, x_3) in the same reference frame). Assume the quantity $\mathbf{r}\left(\mathrm{P}^0, \mathrm{P}\right)$, with physical dimension of length, characterized in the same reference frame by the triplet $(x_1 - x_1^0, x_2 - x_2^0, x_3 - x_3^0)$. If a change of reference frame is made, the new coordinates of the two points, P and P^0, are given by

$$\begin{aligned} X_K &= Q_{Kk} x_k + B_K \\ X_K^0 &= Q_{Kk} x_k^0 + B_K \qquad (K = 1, 2, 3) \end{aligned} \tag{2.59}$$

By subtraction of the two relationships (2.59) is obtained:

$$X_K - X_K^0 = Q_{Kk}(x_k - x_k^0) \qquad (K = 1, 2, 3) \tag{2.60}$$

Relation (2.60), which has the same form as the relationship (2.40), reveals that $\mathbf{r}\left(\mathrm{P}^0, \mathrm{P}\right)$ is a vector. It is called the *position vector* of point P in relation with the point P^0.

Further, it is considered that the coordinates of point P are differentiable functions of a real parameter t: $x_k = x_k(t)$ $(k = 1, 2, 3)$. In this case, a change of reference frame transforms these coordinates into:

$$X_K(t) = Q_{Kk} x_k(t) + B_K \qquad (K = 1, 2, 3) \tag{2.61}$$

The new coordinates are also differentiable functions of the real parameter t. Assume that the point P^0 is fixed. In this case, the velocity vector of the point P in relation with point P^0 can be obtained by using the vector $\mathbf{r}\left(\mathrm{P}^0, \mathrm{P}\right)$, as follows:

$$\mathbf{v}(\mathrm{P}^0, \mathrm{P}) = \frac{d}{dt}\mathbf{r}(\mathrm{P}^0, \mathrm{P}) \tag{2.62}$$

Its physical dimension is length divided by time. The components of this vector are denoted $(\dot{x}_1, \dot{x}_2, \dot{x}_3)$. Using (2.60), it is found that

$$\dot{X}_K(t) = Q_{Kk}\dot{x}_k(t) \qquad (K = 1, 2, 3) \tag{2.63}$$

Relation (2.63) shows that the velocity of a point is a vector, because the components of the velocity of point P do not depend on the coordinates of the point P^0. Thus:

$$\mathbf{v}(\mathrm{P}^0, \mathrm{P}) = \mathbf{v}(\mathrm{P}) \tag{2.64}$$

Assume the case that the functions $x_k = x_k(t)$ $(k = 1, 2, 3)$ have second order derivatives in relation with the parameter t. Repeating the above process, it can be shown that the vector acceleration of the point P, denoted **a**(P), is given by

$$\mathbf{a}(\mathrm{P}) = \frac{d}{dt}\mathbf{v}(\mathrm{P}) \tag{2.65}$$

This vector has physical dimension of length divided by squared time and its components are $(\ddot{x}_1, \ddot{x}_2, \ddot{x}_3)$.

2.7.3.4 The Existence Theorem

Denotes by a_1, a_2, a_3 the triplet of real numbers associated (in a reference frame) with a quantity **a**. In the same reference frame, the components of a vector (or pseudovector) **b** are denoted by b_1, b_2, b_3. The next theorem (given without demonstration) states the necessary and sufficient conditions for a quantity to be vector (or pseudovector).

Theorem 2.4 *(a) If, whatever the vector* ***b****, the sum* $a_k b_k$ *is a scalar (or a pseudoscalar), then* ***a*** *is a vector (or a pseudovector). (b) If, whatever the vector* ***b****, the sum* $a_k b_k$ *is a pseudoscalar (or a scalar) then* ***a*** *is a pseudovector (or a vector).*

2.7.3.5 Change of Components at Reference Frame Transformations

A vector or a pseudovector is completely determined by its components in a certain reference frame. Its components can then be determined in any other reference frame, by the transformation relationships presented below.

The components of the vector **v** can be obtained from relationships (2.57) and (2.58), using the identities (2.47) and (2.53). The procedure is as follows. Multiplying the Eq. (2.58) with Q_{Kj} and by summing up, it is found

$$Q_{Kj}V_K = (\det Q)Q_{Kk}Q_{Kj}v_k = (\det Q)\delta_{kj}v_k = (\det Q)v_j \tag{2.66}$$

Then, taking into consideration the relationship (2.55), it is obtained:

$$v_k = (\det Q)Q_{Kk}V_K \qquad (k = 1, 2, 3) \tag{2.67}$$

Similarly, from the relationship (2.57), it is obtained

$$v_k = Q_{Kk}V_K \qquad (k = 1, 2, 3) \tag{2.68}$$

Relations (2.57) and (2.58) allow the calculation of the new components of the vector, in case that the old components are known. The old components can be calculated as function of the new ones by using the relationships (2.68) and (2.67).

2.7.4 Operations with Scalars and Vectors

The quantities between which a relation exists must have the same physical dimension. The quantities subjected to operations may have the same physical dimension or may have different physical dimensions.

2.7.4.1 Relation of Equality

The relationship of equality makes sense only between quantities with the same physical dimension.

It is said that two scalars (or pseudoscalars) are equal if, in the same reference frame, they are characterized by the same real number. It is said that two vectors (or pseudovectors) are equal if, in the same reference frame, their components are equal.

If two vectors (or pseudovectors) have equal components in a reference frame, then they will have equal components in any other reference frame. To prove this, consider two vectors **u** and **v** which, in two different reference frames, have the components $u_k, v_k \ (k = 1, 2, 3)$ and respectively $U_K, V_K \ (K = 1, 2, 3)$. Using Eq. (2.57), it is found that

$$U_K = Q_{Kk}u_k, \qquad V_K = Q_{Kk}v_k \qquad (K = 1, 2, 3) \tag{2.69}$$

If the two vectors are equal in the first reference frame, then $u_k = v_k \ (k = 1, 2, 3)$. From Eq. (2.69) it is deduced that the equality $U_K = V_K \ (K = 1, 2, 3)$ takes place in the second reference frame.

Similar sentences can be formulated in the case of pseudovectors.

2.7.4.2 Operation of Addition

The addition makes sense only between quantities with the same physical dimension.

The addition of scalars (pseudoscalars) can be done in a given reference frame. It consists in the sum of the numerical values of the scalars (pseudoscalars). Assume two vectors (pseudovectors) **u** and **v**, of components u_k and v_k $(k = 1, 2, 3)$ in a given reference frame. The sum of the two vectors (pseudovectors) is a vector (pseudovector), denoted **w**, whose components are equal to the sum of the numerical values of the components of the two vectors. The addition is written as follows:

$$\mathbf{w} = \mathbf{u} + \mathbf{v} \tag{2.70}$$

To prove that the addition of two vectors (pseudovectors) produces a vector (pseudovector), consider that $w_k = u_k + v_k$ $(k = 1, 2, 3)$ and respectively $W_K = U_K + V_K$ $(K = 1, 2, 3)$, are the components of **w** in two reference frames connected by the relation (2.40). In case of vectors, the transformation relationship (2.57) is applied to the components of the vectors **u** and **v**. It is found that:

$$W_K = U_K + V_K = Q_{Kk}(u_k + v_k) = Q_{Kk}w_k \tag{2.71}$$

In the case that **u** and **v** are pseudovectors, from (2.58) it is obtained

$$W_K = U_K + V_K = (\det Q)Q_{Kk}(u_k + v_k) = (\det Q)Q_{Kk}w_k \tag{2.72}$$

Consequently, by changing the reference frames, the components of the sum of two vectors (pseudovectors) transforms itself as a vector (pseudovector).

2.7.4.3 Scalar Product

The set of all vectors and pseudovectors, regardless of their physical size, will be denoted by V. Assume that **u** and **v** are elements of V. Their components, in a certain reference frame, will be denoted u_k, v_k $(k = 1, 2, 3)$. The *scalar product* (or *inner product*, or *dot product*) of **u** and **v** is defined as the number $u_k v_k$. This product has physical dimension and is denoted $\mathbf{u} \cdot \mathbf{v}$.

If **u** and **v** are vectors, their scalar product is a scalar. This can be easily shown. Denote with U_K and $V_K (K = 1, 2, 3)$ the components of **u** and **v** in a reference frame. By using relations (2.47) and (2.57), it is obtained

$$U_K V_K = Q_{Ki} Q_{Kj} u_i v_j = \delta_{ij} u_i v_j = u_k v_k \tag{2.73}$$

i.e. the quantity $u_k v_k$ is invariant to the transformation (2.40); thus it is a scalar.

In case **u** and **v** are pseudovectors, their scalar product is a scalar. This can be shown using the previous procedure, together with relation (2.58):

$$U_K V_K = (\det Q)^2 Q_{Ki} Q_{Kj} u_i v_j = \delta_{ij} u_i v_j = u_k v_k \tag{2.74}$$

The above results can be applied in the case $\mathbf{v} \equiv \mathbf{u}$. Obviously, the scalar product $\mathbf{u} \cdot \mathbf{u}$ is a scalar, being a quantity invariant to orthogonal transformations. This observation allows the introduction of a quantity which characterizes intrinsically the vector (pseudovector) **u**. This quantity is the module of **u**, defined as $|\mathbf{u}| \equiv (\mathbf{u} \cdot \mathbf{u})^{1/2}$. The module is a scalar with physical dimension equal to the physical dimension of the vector (pseudovector) **u**. The unit vector (or the unit pseudovector) attached to **u** is defined as $\mathbf{w} \equiv \mathbf{u}/|\mathbf{u}|$. This quantity, which has the module equal to the unity, is a dimensionless vector (pseudovector). The vectorial (pseudovectorial) character of this quantity it is easily shown, using relations (2.57) and (2.58):

$$\frac{U_K}{|\mathbf{u}|} = \frac{U_K}{|\mathbf{U}|} = Q_{Kk} \frac{u_k}{|\mathbf{u}|}, \quad \frac{U_K}{|\mathbf{u}|} = \frac{U_K}{|\mathbf{U}|} = (\det Q) Q_{Kk} \frac{u_k}{|\mathbf{u}|} \tag{2.75}$$

The vector (pseudovector) **u** can be represented geometrically through the oriented segment $\mathbf{P'P}$. In this case, $|\mathbf{u}| = |\mathbf{P'P}|$, i.e. the vector's module is numerically equal to the length of the segment $\mathbf{P'P}$. Indeed, using Eq. (2.74), it is obtained by simple processing:

$$|\mathbf{u}| = (u_k u_k)^{1/2} = \left[(x_k - x_k^0)(x_k - x_k^0)\right]^{1/2} = |\mathbf{P'P}| \tag{2.76}$$

This relationship is used to justify why the notion of length of vector (pseudovector) is sometimes used instead of the notion of module.

If **u** is a vector and **v** is a pseudovector, their scalar product is a pseudoscalar. The demonstration involves using relationships (2.57) and (2.58):

$$U_K V_K = (\det Q) Q_{Ki} Q_{Kj} u_i v_j = (\det Q) \delta_{ij} u_i v_j = (\det Q) u_k v_k \tag{2.77}$$

Therefore, the quantity $u_k v_k$ is transformed, indeed, as a pseudoscalar.

2.7.4.4 Vector Product

Assume two elements $\mathbf{u}, \mathbf{v} \in V$. In a reference frame, they have the components u_k and $v_k (k = 1, 2, 3)$, respectively. The *vector product* (or *cross product*) of **u** and **v** is defined as that quantity of components

$$w_i = \in_{ijk} u_j v_k \quad (i = 1, 2, 3) \tag{2.78}$$

Here $\in_{ijk}$ $(i,j,k = 1,2,3)$ is the *symbol of Levi-Civita*, which has only six nonzero components, i.e.:

$$\in_{ijk} = \in_{jki} = \in_{kij} = -\in_{jik} = -\in_{ikj} = -\in_{kji} = 1 \qquad (i,j,k = 1,2,3) \tag{2.78$'$}$$

The vector product is usually written as

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} \tag{2.79}$$

and it has a physical dimension equal to the product of the physical dimensions of **u** and **v**.

Note that if **u** and **v** are vectors, their vector product is a pseudovector. This is easy to prove. The components of **u** and **v** in a different reference frame are denoted by U_K and $V_K (K = 1,2,3)$. Using relations (2.57), the following compact form of the vector product in the same reference frame is found

$$W_I = \in_{IJK} U_j V_K = \in_{IJK} Q_{Jj} Q_{Kk} u_j v_k \tag{2.80}$$

By multiplication with Q_{Ii} and summing, it is obtained

$$Q_{Ii} W_I = \in_{IJK} Q_{Ii} Q_{Jj} Q_{Kk} u_j v_k = \in_{ijk} (\det Q) u_j v_k = (\det Q) w_i \tag{2.81}$$

In other words

$$w_i = (\det Q) Q_{Ii} W_I \tag{2.82}$$

which shows that the result of the vector product is a pseudovector.

The vector product of the pseudovectors **u** and **v** is a pseudovector. The demonstration starts by observing that

$$W_I = \in_{IJK} U_J V_K = \in_{IJK} (\det Q)^2 Q_{Jj} Q_{Kk} u_j v_k = \in_{IJK} Q_{Jj} Q_{Kk} u_j v_k \tag{2.83}$$

The calculus continues in a way similar to the case previously presented.

If one of **u** and **v** is vector and the other one is pseudovector, then their vector product is a vector. The demonstration starts by observing that

$$W_I = \in_{IJK} U_J V_K = \in_{IJK} (\det Q) Q_{Jj} Q_{Kk} u_j v_k \tag{2.84}$$

By multiplication with Q_{Ii} and summing, it is obtained

$$Q_{Ii} W_I = \in_{IJK} (\det Q) Q_{Ii} Q_{Jj} Q_{Kk} u_j v_k = \in_{ijk} (\det Q)^2 u_j v_k = w_i \tag{2.85}$$

which shows that the result of the vector product acts as a vector.

2.7.4.5 Composed Operations

External Product

Assume that S is the set of scalars and pseudoscalars, regardless of their physical dimension. Assume an element $\lambda \in S$ and an element $\mathbf{u} \in V$, having the components $u_k\ (k = 1, 2, 3)$ in a given reference frame. The external product of $\mathbf{u}$ and λ is defined as a quantity having the components $\lambda u_k\ (k = 1, 2, 3)$ in that reference frame. The external product has the physical dimension given by the product of the physical dimensions of $\mathbf{u}$ and λ, and it is denoted as follows:

$$\mathbf{w} = \lambda \mathbf{u} \tag{2.86}$$

The following sentences can be easily verified. If $\mathbf{u}$ is vector and λ is scalar, then $\mathbf{w}$ is vector. If $\mathbf{u}$ is pseudovector and λ is scalar, then $\mathbf{w}$ is pseudovector. If $\mathbf{u}$ is vector and λ is pseudoscalar, then $\mathbf{w}$ is pseudovector. If $\mathbf{u}$ is pseudovector and λ is pseudoscalar, then $\mathbf{w}$ is vector.

Now, it is assumed that λ and μ are scalars with the same physical dimension and $\mathbf{u}$ and $\mathbf{v}$ are vectors with the same physical dimension. In this case, the quantity $(\lambda\,\mathbf{u} + \lambda\,\mathbf{v})$ is a vector. Its physical dimension is different from that of $\mathbf{u}$ and $\mathbf{v}$. Therefore, the external product $\lambda\,\mathbf{u}$ is not closed, i.e. if this operation is applied on a vector $\mathbf{u}$, the result is an element with different physical dimension than that of the vector $\mathbf{u}$. Hence the important conclusion is outlined that, in general, the set of vectors with the same physical dimension, does not form a linear space. In case that K is a field of dimensionless scalars, then the set of vectors with the same physical dimension forms a linear space over K, because the external product is closed.

Mixed Product

Consider three elements $\mathbf{u}, \mathbf{v}, \mathbf{w}$ of V. The *mixed product* (or *scalar triple product*, or *box product*) of $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is defined as being the quantity resulting from the composite operation $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$.

The following sentences be can demonstrated. If $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are vectors, then the mixed product is a pseudoscalar. It can be shown, indeed, that in fact the operation is a scalar product between a vector ($\mathbf{u}$) and a pseudovector ($\mathbf{v} \times \mathbf{w}$). If one of the elements $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ is pseudovector and the other two elements are vectors, their mixed product is a scalar. If two elements of the mixed product are pseudovectors, and the third element is a vector, then their mixed product is a pseudoscalar. If $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are pseudovectors, their mixed product is a scalar.

Vector Triple Product

The vector triple product of three vectors, $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$, is defined as being given by the composite operation $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$.

The following properties can be easily demonstrated. If $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are vectors, then the vector triple product is a vector. If one of $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ is pseudovector and the other two elements are vectors, then the vector triple product is pseudovector. If two elements are pseudovectors and the third is a vector, then the vector triple product is a vector. If $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ are pseudovectors, then the vector triple product is a pseudovector.

References

Beju, I., Soos, E., Teodorescu, P.P.: Tehnici de calcul vectorial cu aplicatii, Editura Tehnica, Bucuresti (1976)

Gellert, W., Kustner, H., Hellwich, M., Kastner, H. (eds): Mica Enciclopedie Matematica, Ed. Tehnica, Bucuresti (1980)

Kaufmann, A., Precigout, M.: Elemente de teoria multimilor si algebra moderna, vol. 2, Ed. Tehnica, Bucuresti (1973)

Mihaileanu, N.: Geometrie analitica, proiectiva si diferentiala, Ed. Didactica si Pedagogica, Bucuresti (1971)

Chapter 3
Vector Calculus and Differential Forms

3.1 Geometrization of Vectors

If the linear algebra is used in physics, the operations used to defining linear and affine spaces must be completed with other types of operations, of less abstract nature. One of these operations consists in the addition (or the composition, in a broader sense) of elements of the same physical quantity, attached to certain points in three-dimensional physical space, having as a result one element of the same physical quantity. It has been already noticed that such a category of elements is that of the physical vectors. In the following, the theory of the physical vectors will be complemented with other aspects (Beju et al. 1976).

Physical vectors are usually represented in a space with three dimensions (in classical physics) or in a four-dimensional space (in case of the theory of relativity). There are, however, some categories of physical vectors that require representations in n-dimensional spaces. In this chapter, the standard classification of the vectors will be presented, with exemplification in the three-dimensional space. Choosing the three-dimensional space has the advantage that it attaches to a vector the intuitive geometric image of an oriented segment.

3.1.1 Types of Vectors

In classical physics there are three main types of vectors: *bound vectors*, *sliding vectors* and *free vectors*. These types are defined below and their properties are briefly presented.

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_3

3.1.1.1 Three-Dimensional Reference Frames

Assume a three-dimensional Euclidean space, denoted E_3. A reference frame in this space consists of three orthogonal axes $Ox_i (i = 1, 2, 3)$ that intersect at a point denoted O, called the *origin* of the reference frame. A reference frame is called right-handed if, when looking in the positive direction of the axis Ox_i, the overlap of the axis Ox_j on the axis Ox_k can be achieved through a rotation angle $\pi/2$ in the positive sense (from right to left). For a *right-handed reference frame* the following convention is adopted:

$$(i, j, k) = (1, 2, 3) \tag{3.1}$$

that means that the indices i, j, k of the coordinate axes take the distinct values $1, 2, 3$ (in that order), or after a cyclic permutation.

3.1.1.2 Oriented Segments (Bound Vectors)

Assume a reference frame in the space E_3 and two points $\mathrm{A}(x_1, x_2, x_3)$ and $\mathrm{B}(y_1, y_2, y_3)$ in the same space. The ordered pair of points $(\mathrm{A,B})$ is called *oriented segment* (or *bound vector*). The point A is called the origin of the segment and the point B is called the segment end or endpoint. The oriented segment will be denoted **AB**. The numbers $u_i = y_i - x_i (i = 1, 2, 3)$ are called the *canonic coordinates* of the oriented segment. The set of oriented segments (bound vectors) will be denoted V_l.

3.1.1.3 Equipollent Oriented Segments (Bound Vectors)

It is said that the oriented segments are equipollent if they have the same canonical coordinates. Assume two points, $\mathrm{A}'(x'_1, x'_2, x'_3)$ and $\mathrm{B}'(y'_1, y'_2, y'_3)$. They are chosen in such way that to lead to an oriented segment $\mathbf{A}'\mathbf{B}'$, equipollent with **AB**. According to the definition of equipollence, the following relationships take place: $y_i - x_i = y'_i - x'_i (i = 1, 2, 3)$. These relationships are equivalent with $(x'_i + y_i)/2 = (x_i + y')/2$, which means that the quadrilateral $\mathrm{ABB}'\mathrm{A}'$ is a parallelogram. The relation of equipollence is an equivalence relationship defined on the set of bound vectors V_l. Denote with V the set of equivalence classes defined by this equivalence relation. A certain element $\mathbf{u} \in V$ is characterized only by its canonical coordinates (u_1, u_2, u_3) associated with the equivalence class $\mathbf{u}$. Obviously, all oriented segments of the class $\mathbf{u}$ have the same canonical coordinates.

An operation of addition-like can be defined between two elements $\mathbf{u}, \mathbf{v} \in V$, the result of the operation being the class characterized by the sum of the canonical coordinates of $\mathbf{u}$ and $\mathbf{v}$. An operation of product-like can be also defined between a class (u_1, u_2, u_3) and a scalar α, the result of the operation being the class characterized by the canonical coordinates $(\alpha u_1, \alpha u_2, \alpha u_3)$. With these two operations, it becomes immediately apparent that V/R has the structure of a three-dimensional real linear space.

The oriented segment **AB**, with canonical coordinates (u_1, u_2, u_3), will be assigned as a representative of the class $\mathbf{u} \in V$. Similarly, the oriented segment **BC**, of canonical coordinates (v_1, v_2, v_3), will be designated as a representative of the class $\mathbf{v} \in V$. The following relationships take place between the canonical coordinates of **u** and **v** and the coordinates of three points, $\mathrm{A}(\mathrm{x}_1, x_2, x_3)$, $\mathrm{B}(\mathrm{y}_1, y_2, y_3)$ and $\mathrm{C}(\mathrm{z}_1, z_2, z_3)$: $y_i - x_i = u_i, z_i - y_i = v_i$ (i = 1, 2, 3). By adding the two relationships one finds $z_i - x_i = u_i + v_i$. But the left member of this equality represents the canonical coordinates of the oriented segment **AC**. Therefore, this segment will be representative of the addition $\mathbf{w} = \mathbf{u} + \mathbf{v}$, since the right member of equality corresponds to this addition. The addition in V is carried out by means of the representatives, using the so-called *triangle rule (*or *parallelogram rule).*

3.1.1.4 Characteristics of Bound Vectors

By definition, for full determination of an oriented segment (bound vector) should be specified its origin A (also known as point of application) and its endpoint B. Once these two points are specified, one can determine the size of the bound vector (oriented segment) as the length of the segment **AB**, i.e. the distance between the points A and B. Sometimes, instead of *magnitude (or length) of a bound vector* it is used the denomination *module of a bound vector*. The support of the bound vector is the straight line determined by the points A and B. Also, the direction of the bound vector is the direction of its support. The order in which the two points appear in the ordered pair (A,B) specifies the sense of the bound vector **AB**. Therefore, once the direction of a vector has been fixed (by choosing the sense of the bound vector) the application point and the endpoint of the vector have been also determined.

If a bound vector is assumed as being known, all its features are also assumed as being known: the module, the support, the direction, the origin and the endpoint. But these features are not independent each other. To determine a bound vector starting from the knowledge of its characteristics it is sufficient to specify:

(a) the endpoint, the module, the direction and the sense or
(b) the application point and its endpoint; or
(c) the application point, the module, the direction and the sense.

3.1.1.5 Sliding Vectors

Assume two bound vectors **AB** and $\mathbf{A'B'}$. They are called equivalent vectors if the following two conditions are fulfilled:

1. the points A,B,A', B' are collinear and
2. the bound vectors have the same length and the same sense.

Note that, in order for two bound vectors to be equivalent, it is not necessary for them to have the same origin or endpoint. One cans easy show that the properties 1 and 2 induce an equivalence relation on the set V_l. Denote by V_a the set of the equivalence classes associated with this equivalence relation. In this case, an element u $\in V_a$ will be characterized only by its module, support and sense and will be called *sliding vector*. Naturally, V_a is called the set of sliding vectors.

3.1.1.6 Free Vectors

A free vector is a mathematical object characterized by direction, sense and module. If the free vector is denoted by **V**, then its module is denoted |**V**| or simply V. The following property is fulfilled: $\mathrm{V} \geq 0$. The free vector is completely characterized by its canonical coordinates $\mathrm{V_i}(i = 1, 2, 3)$, which constitute an ordered triplet of numbers. The module of the vector can be calculated by using the canonical coordinates as follows:

$$\mathrm{V} = (V_i V_i)^{1/2} = (V_1^2 + V_2^2 + V_3^2)^{1/2} \tag{3.2}$$

The relation (3.2) has a simple geometric interpretation: it allows calculating the length of the diagonal of a right parallelepiped as a function of the length of its sides.

Note that in the first equality in (3.2) the *Einstein summation convention* has been used. That convention states that when an index variable appears twice in a single term and is not otherwise defined (such an index is called *dummy index*), it implies summation of that term over all the values of the index. The same convention defines a *free index* as that index which appears only once in monomial and states that an index cannot appear three times in a monomial.

3.1.1.7 Operations with Vectors

Defining the operations with vectors can be done in two ways. First, the features introduced above can be used: module, direction, sense, support. This procedure is often used in sciences, with the benefit of a precise physical significance. Defining the vector operations can be also done in a more abstract way, purely algebraically, by using the canonical coordinates.

3.1.1.8 Equality of Two Vectors

The *relation of equality* can be defined on the set of vectors of the same type (bound vectors, sliding vectors, free vectors, respectively). It is said that two vectors $\mathbf{V}_1$ and

$\mathbf{V}_2$ of the same type are equal if the elements that characterize them are equal each other. First, assume the case of the free vectors $\mathbf{V}_1(V_{1i})$ and $\mathbf{V}_2(V_{2i})$. The equality of these vectors is denoted

$$\mathbf{V}_1 = \mathbf{V}_2 \tag{3.3}$$

or, by using the components:

$$V_{1i} = V_{2i} \quad (i = 1, 2, 3) \tag{3.4}$$

If the equality refers to bound vectors, in addition to the condition (3.3), the vectors must also have the same application point. In the case of the equality of two sliding vectors, the condition of having equal free vectors is supplemented by the condition that the two vectors must have the same support.

Assume the vectors $\mathbf{V}_1$, $\mathbf{V}_2$ and $\mathbf{V}_3$. The relation of equality between these vectors has the following properties:

(a) $\mathbf{V}_1 = \mathbf{V}_1$ (the relation is *reflexive*)
(b) $\mathbf{V}_1 = \mathbf{V}_2 \Leftrightarrow \mathbf{V}_2 = \mathbf{V}_1$ (the relation is *symmetric*)
(c) $\mathbf{V}_1 = \mathbf{V}_2, \mathbf{V}_2 = \mathbf{V}_3 \Leftrightarrow \mathbf{V}_1 = \mathbf{V}_3$ (the relation is *transitive*)

These properties are valid for any type of vectors.

3.1.1.9 Geometric Characterization of the Types of Vectors

The types of vectors defined above are characterized by a number of common geometric properties, as well as by other geometrical properties specific for each type.

Bound Vectors

Remember that the bound vector is a mathematical object characterized by direction, sense, module and origin (point of application). Assume a coordinate system Ox_i. The position vector of an arbitrary point A, denoted $\mathbf{r}(x_1, x_2, x_3)$, has the application point in the origin O of the coordinate system and the end point in the point A. It is a bound vector. An arbitrary bound vector V has the origin in a certain point (denoted, for example, by A) which is specified by the vector position $\mathbf{r}$. The terminal point of the vector $\mathbf{V}$ is found in a different point, denoted for example B. Therefore, the bound vector $\mathbf{V} = \mathbf{AB}$ (where $\mathbf{AB}$ is an oriented segment) may be characterized by the bound vector $\mathbf{r}$ and the free vector $\mathbf{V}$ (its representative in the point A). In other words, the bound vector is characterized by two ordered triplets of numbers (x_i and $V_i (i = 1, 2, 3)$, respectively).

Sliding Vectors

The sliding vector is a mathematical object that can be characterized by direction, sense, module and support. The vector is placed on a specified straight line but its application point is not specified. Hence the common assertion that such a vector can slide on the support line.

Similarly to a bound vector, the sliding vector can be specified by two triplets of numbers (x_i and $V_i(i = 1, 2, 3)$, respectively). However, in case of a sliding vector a certain relationship exists between these two triplets. Therefore, a sliding vector is specified by five real numbers. To obtain a bound vector from a sliding vector, knowledge of a real number is also needed; this allows the specification of the application point of the bound vector on the support line, with reference to an origin chosen on that line.

Operations with sliding vectors are performed by using representatives of these vectors.

Free Vectors

In the case that two equal free vectors are bounded to two different application points, two bound vectors are obtained, which are equipollent. Note that a bound vector can be considered as an element of the set of equipollent vectors that have the same direction, the same sense and the same module.

Operations with free vector are performed by using representatives of these vectors.

3.1.1.10 Algebraic Structures on Sets of Vectors

Note that, in general, the set of bound vectors V_l and the set of sliding vectors V_a do not form a group. Therefore, for these sets one cannot speak of a linear space structure. This distinguishes the two sets of vectors from the set V/R of equipollent vectors, which has the structure of a linear space.

3.1.2 Geometrization of Physical Vectors

In order to represent a vector as an oriented segment, a reference frame $ox_1x_2x_3$ will be considered. Also, consider the vector (or pseudovector) **w**, which has components (w_1, w_2, w_3) in that reference frame. Two points must be identified, $\mathrm{P}^0(x_1^0, x_2^0, x_3^0)$ and $\mathrm{P}(x_1, x_2, x_3)$, respectively, in such a way that

$$x_k - x_k^0 = w_k \qquad (k = 1, 2, 3) \tag{3.5}$$

Relations (3.5) can be seen as a system of equations in the unknowns $x_k, x_k^0 (k = 1, 2, 3)$. The matrix of the coefficients of this system has the form:

$$\left\{\begin{matrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{matrix}\right\} \tag{3.6}$$

This is a matrix of rank three. The system of equations is compatible and indeterminate. Consider $x_k (k = 1, 2, 3)$ as the main unknowns, the other unknowns being chosen as independent parameters. This is equivalent to choose an arbitrary point P^0. In this case, the point P is uniquely determined by solving the system (3.5).

What is essential in this procedure is that a vector (or pseudovector) was associated with a set of segments P^0P, which is obtained by solving the Eq. (3.5). It can be said that these segments are oriented, because the meaning of the point P^0 is different from the meaning of the point P (the first point is the origin and the second is the end of the segment, respectively).

Note that the vectors which are associated with oriented segments lose their own physical dimension, by obtaining the dimension of length.

3.1.2.1 Geometrization of Free Physical Vectors

Assume that the vector (or pseudovector) **w** is not associated with a certain point, being a free vector (pseudovector). In this case, the vector can be associated with all the set of oriented segments. Consider two oriented segments, labeled $\mathbf{P^0P}$ and $\mathbf{P'^0P'}$. The coordinates of the four points that define the segments fulfill the Eq. (3.5), i.e.

$$x'_k - x'^0_k = x_k - x^0_k \qquad (k = 1, 2, 3) \tag{3.7}$$

This is equivalent to saying that the segments $\mathbf{P^0P}$ and $\mathbf{P'^0P'}$ are parallel and equal, or to stating that the quadrilateral $P^0PP'P'^0$ is a parallelogram.

3.1.2.2 Geometrization of Bound Physical Vectors

In physics there are numerous examples in which the vector (or pseudovector) **w** is bound, being attached to a specific point in space (e.g. the velocity vector or the acceleration vector). In this case, that point will be chose the origin P^0 of the vector. By solving the system (3.5), the point P is obtained in a unique way.

3.1.2.3 Geometrization of Sliding Physical Vectors

In some branches of physics (e.g. the solid mechanics) the vector (or pseudovector) **w** is sliding, i.e. it is associated with a straight line whose vector parameters are precisely the components of that vector. In this situation, it is said that the point P^0 can "slide" on the line determined by P and P^0. The components of the vectors (pseudovectors) get a clear geometric meaning through this procedure; they represent the projections on the coordinate axes of the oriented segment $\mathbf{P^0P}$.

3.1.3 Representations of Vectors in Given Bases

The elementary results presented below refer to the specific case of the real space R^3. They can however be easily generalized to the case of n-dimensional spaces.

3.1.3.1 The Contravariant Components of a Vector

A positive basis is considered, consisting of the vectors $\mathbf{e}_i (i = 1, 2, 3)$. By definition, such a basis fulfills the condition that the mixed product of the base vectors is a positive number:

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) > 0 \tag{3.8}$$

Consider a certain vector $\mathbf{V}$. Regardless of its type (free vector, bound vector or sliding vector) this vector can be represented as

$$\mathbf{V} = V^1\mathbf{e}_1 + V^2\mathbf{e}_2 + V^3\mathbf{e}_3 \equiv V^i\mathbf{e}_i \tag{3.9}$$

Here $V^i (i = 1, 2, 3)$ are the components of the vector $\mathbf{V}$ in the basis $\{\mathbf{e}_i\}$ and in the last equality of (3.9) the Einstein summation convention has been used (which will also be used further). The components $\mathbf{V}^i$ are called *contravariant*. They can be expressed as (Beju et al. 1976, p. 41):

$$V^i = \frac{\in^{ijk} (\mathbf{V}, \mathbf{e}_j, \mathbf{e}_k)}{2(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)} \quad (i = 1, 2, 3) \tag{3.10}$$

where the *Levi-Civita permutation symbol* has been used. Note that relations (3.10) are similar with the *Cramer rule* in the theory of linear algebraic equations. This observation allows the introduction of a *reciprocal basis* (or *dual basis*), defined by using the vectors

$$\mathbf{e}^i = \frac{\in^{ijk} \mathbf{e}_j \times \mathbf{e}_k}{2(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)} \quad (i = 1, 2, 3) \tag{3.11}$$

These vectors are normal on the planes formed with the vectors of the initial basis, taken in pairs. It is noted that if the basis vectors are *polar vectors* (or, in other words, free vectors, bound vectors or sliding vectors) then the vectors of the reciprocal basis are *axial vectors* (or, in other words, pseudovectors).

Another relationship that occurs is:

$$\mathbf{e}_i = \frac{\in_{ijk} \mathbf{e}^j \times \mathbf{e}^k}{2(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3)} \quad (i = 1, 2, 3) \tag{3.12}$$

which means that the reciprocal of the reciprocal basis is the initial basis. One can notice that:

$$(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3) = \frac{1}{(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)} > 0 \tag{3.13}$$

Another way to write relations (3.10) is based on using the scalar product, as follows:

$$V^i = \mathbf{V} \cdot \mathbf{e}^i \quad (i = 1, 2, 3) \tag{3.14}$$

3.1.3.2 The Covariant Components of a Vector

Regardless of its type (bound vector, sliding vector or free vector), the vector **V** can be represented in the reciprocal basis $\{\mathbf{e}_i\}$ by using the relationship

$$\mathbf{V} = V_i \mathbf{e}^i \tag{3.15}$$

Here $V_i(i = 1, 2, 3)$ are the *covariant components* of the vector. The covariant components allow to writing relations which are similar to those written above by using the contravariant components. Thus, these components may be obtained from the formulas

$$V_i = \frac{\in_{ijk} (\mathbf{V}, \mathbf{e}^j, \mathbf{e}^k)}{2(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3)} \quad (i = 1, 2, 3) \tag{3.16}$$

or may be written by using the scalar product, as

$$V_i = \mathbf{V} \cdot \mathbf{e}_i \quad (i = 1, 2, 3) \tag{3.17}$$

3.1.3.3 Transition from One Type of Component to Another

The following relationship, involving both the vectors of the base and the vectors of the reciprocal base, takes place:

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta^i_j \quad (i, j = 1, 2, 3) \tag{3.18}$$

In the relation (3.18) the Kronecker's symbol δ^j_i has been used, defined in a way similar to the definition of δ_{ij} given in Eq. (2.46), i.e.:

$$\delta_i^j \equiv \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \tag{3.19}$$

Relations presented in previous subsections allow to writing the vector $\mathbf{V}$ under one of the following forms

$$\mathbf{V} = (\mathbf{V} \cdot \mathbf{e}_j)\mathbf{e}^j = (\mathbf{V} \cdot \mathbf{e}^j)\mathbf{e}_j \tag{3.20}$$

Next, the so-called *coefficients of the fundamental form,* denoted g_{ij} and g^{ij}, respectively, are defined. The definitions are:

$$\begin{aligned} &g_{ij} \equiv \mathbf{e}_i \cdot \mathbf{e}_j \\ &g^{ij} \equiv \mathbf{e}^i \cdot \mathbf{e}^j \quad (i,j = 1,2,3) \end{aligned} \tag{3.21, 22}$$

Relationships (3.20) and the definitions (3.21) and (3.22) are now used for two specific cases, namely if $\mathbf{V} = \mathbf{e}_i$ and $\mathbf{V} = \mathbf{e}^i$, respectively, yielding the relationships:

$$\mathbf{e}_i = g_{ij}\mathbf{e}^j, \quad \mathbf{e}_i = g^{ij}\mathbf{e}_j \quad (i = 1,2,3) \tag{3.23}$$

allowing the calculation of the vectors of the initial base $\{\mathbf{e}_i\}$ as a function of the vectors of the reciprocal basis $\{\mathbf{e}^j\}$, and vice versa. The system consisting of the first linear Eq. (3.23) can be solved for finding the unknown $\{\mathbf{e}_i\}$ as function of the vectors $\{\mathbf{e}^j\}$, only if this system is nonsingular. This implies that the determinant g of the system, which is given by

$$g \equiv \det|g_{ij}| = \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} \tag{3.24}$$

must be different from zero. It can be shown that (Beju et al. 1976, p. 43)

$$g = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)^2 > 0 \tag{3.25}$$

The normalized minor of the element g_{ij} of the determinant of (3.24) is denoted by g^{ji}. Because of the relation of symmetry $g_{ij} = g_{ji}$, which is obvious considering the definition (3.21), the following result is obtained: $g^{ji} = g^{ij}$. It is noted that the normalized minor corresponds to the relations (3.22) and the following relationships take place:

$$g_{ik}g^{kj} = \delta_i^j, \quad g^{ik}g_{kj} = \delta_j^i \qquad (i,j = 1,2,3) \tag{3.26}$$

Taking into account the relationship (3.13), the following result is obtained:

$$\det|g^{ij}| = \frac{1}{g} \tag{3.27}$$

The scalar multiplication of the relations (3.23) by **V** leads to the following formulas

$$V_i = g_{ij}V^j \quad V^i = g^{ij}V_j \qquad (i = 1, 2, 3) \tag{3.28}$$

These formulas allow the calculation of the contravariant components of a vector in function of the covariates components, and vice versa.

3.1.3.4 Vectors as Dual Quantities

Relationships (3.28) allow the observation that a vector can be represented in two forms, characterized by contravariant components and by covariant components, respectively. It is said that the vector is a *dual quantity*. An important consequence is that the module of a vector can be expressed by any of the relationships

$$\mathbf{V}^2 = g_{ij}V^iV^j = g^{ij}V_iV_j = V_iV^i \tag{3.29}$$

3.1.3.5 Comments on Bases

Orthonormal Basis

In the particular case of an orthonormal basis, the following relationships take place:

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = (\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3) = 1 \tag{3.30}$$

Also, the reciprocal basis coincides with the given basis and the coefficients of the fundamental form g_{ij}, as well as the coefficients g^{ij}, become Kronecker symbols. Consequently, the covariant components coincide with the contravariant components. In this case it is customary to use only subscripts. Also, one uses a single set of unit vectors (sometimes denoted by $\{\mathbf{i}_j(j = 1, 2, 3)\}$), which specify the given orthonormal basis.

Orthogonal Basis

If the basis is orthogonal, the coefficients g_{ij} and g^{ij} are zero for $i \neq j$, but they are not Kronecker symbols, because $g_{11}, g_{22}, g_{33}, g^{11}, g^{22}, g^{33}$ are generally not equal to unity.

Negative Basis

Instead of a positive base one can also use a negative base. In this situation the reciprocal basis is also negative.

3.2 Elements of Vector Calculus

3.2.1 Scalar and Vector Fields

An excellent guide to vector calculus is the work by Beju et al. (1976), which will be followed closely. A scalar field is defined by using a map that transforms every point P of the physical space, characterized by the position vector $\mathbf{r}(x_1, x_2, x_3)$, into a real number. The map is denoted $(x_1, x_2, x_3) \to U(x_1, x_2, x_3)$. Assume that the function $U = U(\mathrm{P})$ has continuous derivatives of the first order. In the next sections the need for higher order derivatives will be mentioned explicitly.

A vector field is defined by using a map that transforms every point P of the physical space into a vector. The map is denoted $(x_1, x_2, x_3) \to \mathbf{V}(x_1, x_2, x_3)$. The vector which is involved in this definition is a bound vector, having the point P as the application point. The function $\mathbf{V} = \mathbf{V}(\mathrm{P})$ has continuous first order derivatives.

After choosing the basis $\left\{\mathbf{i}_j (j = 1, 2, 3)\right\}$, the vector $\mathbf{V} = \mathbf{V}(x_1, x_2, x_3) = \mathbf{V}(\mathbf{r})$ can be written as:

$$\mathbf{V}(x_1, x_2, x_3) = V_j(x_1, x_2, x_3)\mathbf{i}_j \tag{3.31}$$

where $V_j (j = 1, 2, 3)$ are the components of the vector. For further developments, it is useful to introduce some vector fields, which involve spatial derivatives of the components of the vector:

$$\frac{\partial \mathbf{V}}{\partial x_i} = \mathbf{V}_{,i} = \frac{\partial \mathbf{V}_j}{\partial x_i}\mathbf{i}_j = \mathbf{V}_{j,i}\mathbf{i}_j \tag{3.32}$$

and to introduce the differential of the vector, defined as follows:

$$d\mathbf{V} = \mathbf{V}_{,i} dx_i \tag{3.33}$$

In previous relationships, the subscript which appears to the right of the comma shows the variable for which the differentiation is performed.

3.2.1.1 Lines of Vectors

The lines of vectors are curves with the property that the tangents at any point P have the direction $\mathbf{V}(\mathbf{r})$. Lines of vectors form a congruency (or, in other words, a family) of curves. One can show that the differential d$\mathbf{r}$ of the position vector is tangent to the lines of vectors. As a result, the vector equation of the lines of vectors has the form

$$\mathbf{V}(\mathbf{r}) \times d\mathbf{r} = 0 \tag{3.34}$$

Lines of vectors can be determined by solving the following system of first order differential equations:

$$\frac{dx_1}{V_1} = \frac{dx_2}{V_2} = \frac{dx_3}{V_3} \tag{3.35}$$

3.2.1.2 Potential

It is considered a scalar field U and a field vector **V**, defined by relations

$$V_i = U_{,i} \quad (i = 1, 2, 3) \tag{3.36}$$

It is also considered a certain unit vector $\mathbf{n}(n_1, n_2, n_3)$. Notice the following relationship:

$$\mathbf{V} \cdot \mathbf{n} = U_{,i} n_i = \frac{\partial U}{\partial \mathbf{n}} \tag{3.37}$$

On a new three-orthogonal reference frame, $Ox'_1 x'_2 x'_3$, the components of the vector **V** defined by (3.36) are $\partial U/\partial x'_i (i = 1, 2, 3)$. This shows that the definition of the vector field **V** does not depend on the chosen reference frame. The vector field defined by (3.36) is called *conservative field* and the appropriate vectors are called *conservative vectors*. Also, it is said that a conservative field derives from the *potential field* (or, in short, from *the potential*) U.

To obtain the necessary and sufficient conditions for the scalar functions

$$V_i = V_i(x_1, x_2, x_3; t) \quad (i = 1, 2, 3) \tag{3.38}$$

to be components of a quasi-conservative vector, the relationships (3.36) are used. Also, it is required that the mixed second order derivatives of the quasi-potential $U = U(x_1, x_2, x_3; t)$ (for which it is presumed that it has second order derivatives with respect to spatial variables) be independent of the differentiation order. The following conditions are obtained:

$$\in_{ijk} \partial_j V_k = \in_{ijk} V_{k,j} = 0 \quad (i = 1, 2, 3) \tag{3.39}$$

Equipotential Surfaces

Assume a surface with the property that in any point the scalar potential has a constant value. The equation of this surface has the form:

$$U(x_1, x_2, x_3; t) = C(= \text{const}) \tag{3.40}$$

Assume that U is a differentiable function. The surface defined by the Eq. (3.40) is called equipotential surface. If the potential depends on the spatial coordinates and a parameter t, i.e. $U = U(x_1, x_2, x_3; t)$, it is said that it is an *equi-quasi-potential surface*. It is written:

$$U(x_1, x_2, x_3; t) = C(= \text{const}) \tag{3.41}$$

In case that the point (x_1^0, x_2^0, x_3^0) belongs to two equipotential surfaces, the following relations can be written:

$$U(x_1^0, x_2^0, x_3^0) = C_1, \quad U(x_1^0, x_2^0, x_3^0) = C_2 \tag{3.42}$$

where C_1 and C_2 are constants. By subtraction of the above relationships it is found that $C_1 = C_2$. Since the scalar function is assumed to be uniform, it is deduced that these two equipotential surfaces coincide. Another consequence is that if two equipotential surfaces do not coincide, they have no common point.

Gradient

The following vector differential operator is due to Hamilton:

$$\nabla \equiv \mathbf{i}_j \frac{\partial}{\partial x_j} = \mathbf{i}_j \partial_j \tag{3.43}$$

It is called *gradient*, or *Hamilton operator*, or *"nabla" operator.* With this operator, the vector field $\mathbf{V}$ can be expressed as a function of the potential U, under the form:

$$\mathbf{V} = U_{,j}\mathbf{i}_j = \left(\mathbf{i}_j \frac{\partial}{\partial x_i}\right) U = \nabla U \tag{3.44}$$

If the operator ∇ is applied on the scalar function U, a vector function called the gradient of U is obtained, which is denoted as follows:

$$\operatorname{grad} U = \nabla U \tag{3.45}$$

Therefore, the operator ∇ transforms a scalar field into a vector field. The gradient provides information about the variation of a scalar function. Knowing the gradient is equivalent to the knowledge of the first order derivatives of the scalar function. Using the gradient definition, the Eq. (3.37) can be rewritten as

$$\mathbf{n} \cdot \operatorname{grad} U = \frac{\partial U}{\partial n} \tag{3.46}$$

where $\partial U/\partial n$ is the derivative of the scalar field U on the direction of the unit vector n. Therefore, the gradient forms a vector field, regardless of the chosen reference frame.

Gradient Properties

The gradient has a number of properties, whose knowledge often simplify the calculation. Consider two scalar functions U_1 and U_2 and a differentiable and integrable function f. A list of gradient properties follows. They can be used for potential fields as well as for quasi-potential fields:

$$\begin{aligned}
&\operatorname{grad}(U_1 + U_2) = \operatorname{grad} U_1 + \operatorname{grad} U_2 \\
&\operatorname{grad}(U_1 U_2) = U_1 \operatorname{grad} U_2 + U_2 \operatorname{grad} U_1 \\
&\operatorname{grad}(CU) = C \operatorname{grad} U, \quad (C = \text{const}) \\
&\operatorname{grad} C = 0, \quad (C = \text{const}) \\
&\operatorname{grad} f(U) = f'(U) \operatorname{grad} U \\
&\operatorname{grad} f(U_1, U_2) = f'_{U_1} \operatorname{grad} U_1 + f'_{U_2} \operatorname{grad} U_2 \\
&f(U) \operatorname{grad} U = \operatorname{grad} \int f(U) dU
\end{aligned} \tag{3.47a–g}$$

For the particular case of the position vector $\mathbf{r}$, the following properties apply:

$$\begin{aligned}
&\operatorname{grad} \mathbf{r} = \frac{1}{r} \mathbf{r}, \\
&\operatorname{grad} f(r) = \frac{f'(r)}{r} \mathbf{r}, \\
&\operatorname{grad}(\mathbf{C} \cdot \mathbf{r}) = \mathbf{C}, \quad \mathbf{C} = \mathbf{const}.
\end{aligned} \tag{3.48a–c}$$

Differential and Derivative of a Scalar Potential

If the field potential depends on a parameter, i.e. $U = U(\mathbf{r}; t)$, a vector field can be defined by using relations (3.36). It is a quasi-conservative field, its vectors being quasi-conservative vectors. Similarly, the scalar function U is called quasipotential. One finds that a conservative vector field is a gradient field.

The differential of a quasipotential can be expressed as follows

$$dU = U_{,j} dx_j + \dot{U} dt = \nabla U \cdot d\mathbf{r} + \dot{U} dt \tag{3.49}$$

where $\dot{U} \equiv \partial U/\partial t$. Consider the case when the position vector depends on the parameter t, i.e. $\mathbf{r} = \mathbf{r}(t)$. The *total derivative* (also called the *substantial derivative*) of the quasipotential is given by the sum of the spatial derivative and temporal derivative, i.e.:

$$\frac{dU}{dt} = U_{,j} \dot{x}_j + \dot{U} = \nabla U \cdot \dot{\mathbf{r}} + \dot{U} \tag{3.50}$$

In case of a potential, the relation (3.50) is reduced only to the spatial derivative.

Consider the potential $U = U(x_1, x_2, x_3) = U(\mathbf{r})$. The relationships (3.49) and (3.50) leads to the following form of its differential dU:

$$dU = \operatorname{grad} U \cdot d\mathbf{r} \tag{3.51}$$

Relation (3.51) proves that the symbol d of the total differential can be conceived as an operator, expressed by the following scalar product

$$d = d\mathbf{r} \cdot \nabla = d\mathbf{r} \cdot \operatorname{grad} \tag{3.52}$$

Knowledge of a quasi-conservative field, allows obtaining the scalar quasipotential from which it is derived, by using a curvilinear integral on an arbitrary curve:

$$\int\limits_{P_0P} V_j dx_j \tag{3.53}$$

Here, the point P$_0$ is considered fixed and the point P is considered variable. First, the variable t is assumed as a parameter. In this case, the scalar quasipotential is reduced to the form $U = U(x_1, x_2, x_3)$. The integrand in (3.53) becomes an exact total differential, since the curvilinear integral (3.53) does not depend on the path. The properties of the quasipotential do not change by adding an arbitrary function of variable t. Similarly, the properties of a potential $U = U(x_1, x_2, x_3)$ do not change by adding an arbitrary constant.

Gradient and Equipotential Surfaces

Consider a curve C, on which a point $P(x_1, x_2, x_3)$ is moving. Assume that the differential $d\mathbf{r}$ of the position vector $\mathbf{r}$ has the direction of the tangent to the curve C. By canceling the expression (3.49) it follows that, in case of a scalar quasipotential field, the C curve cannot be found on a corresponding equi-quasipotential surface. In case of a potential field, the curve C can be found on a corresponding equipotential surface.

The gradient can be written as a function of the unit vectors of the coordinate system, as follows:

$$\operatorname{grad} U = U_{,j}\mathbf{i}_j \tag{3.54}$$

Therefore, the gradient of the scalar function U is normal to the equipotential surface $U(x_1, x_2, x_3) - C = 0$. Next, define $\mathbf{n}$ as the unit vector of $\operatorname{grad} U$. Using Eq. (3.54), it follows that the sense of the gradient of U is in such a way that the value of U increases. Therefore, the congruence of the gradient lines is normal to the family of the corresponding equipotential surfaces, their sense being that in which the value of U increases. One may check that these properties remain valid for the particular case of quasi-conservative scalar fields.

Work of a Conservative Vector

A quantity often encountered in mechanics and thermodynamics is *work*. For the particular case of a field of conservative vectors, the elementary work has the expression

$$dL = \nabla U \cdot d\mathbf{r} = dU \tag{3.55}$$

Note that the above expression is an exact total differential. Therefore, by integrating on a finite distance between the points P_0 and P_1 one finds:

$$L_{\mathrm{P}_0\mathrm{P}_1}(V) = U(\mathrm{P}_1) - U(\mathrm{P}_0) \tag{3.56}$$

It is found that, in the case of a conservative vector, the work does not depend on the path between two points but only on the potential values in those points.

The *circulation of a vector* on a closed curve is defined as the work of the vector on that closed curve. One may prove that the circulation of a conservative vector on a closed curve is zero.

Directional Derivative

A curve C can be defined by a function $s \to \mathbf{r}(s)$ that applies the set of real numbers into the set of position vectors. The parameter s may optionally be identified with the time. A scalar potential $U(\mathbf{r})$ associated with the respective curve becomes $U = U(s)$. This leads to:

$$\frac{\partial U}{\partial s} = \frac{dU}{ds} = \operatorname{grad} U \cdot \frac{d\mathbf{r}}{ds} = \operatorname{grad} U \cdot \boldsymbol{\tau} \tag{3.57}$$

This is the derivative on the direction of unit vector $\boldsymbol{\tau}$ of the tangent to the curve C. Note that in this case the partial derivative is equal to the total derivative.

Similarly, one can introduce the operator *directional derivative* on the direction of the unit vector $\mathbf{n}$. For this, use relation (3.46) and obtain:

$$\frac{\partial}{\partial n} \equiv n \cdot \nabla = n \cdot \operatorname{grad} \tag{3.58}$$

For the particular situation in which $\mathbf{r} = \mathbf{r}(s)$ and $\mathbf{n} = \boldsymbol{\tau}$, one obtains the operator

$$\frac{\partial}{\partial s} = \frac{d}{ds} = \mathbf{r}'(s) \cdot \nabla = \boldsymbol{\tau} \cdot \operatorname{grad} \tag{3.59}$$

Differential and Derivative of a Vector Field

Assume a vector $\mathbf{A}$, constant or variable. Using this vector, the following linear scalar differential operator $\mathbf{A} \cdot \nabla$ can be defined:

$$\mathbf{A} \cdot \nabla = \mathbf{A} \cdot \operatorname{grad} = A_i \frac{\partial}{\partial x_i} = A_i \partial_i \tag{3.60}$$

This operator can be applied to the scalar field U. The result is the scalar

$$(\mathbf{A} \cdot \mathrm{grad})U = (A_i \partial_i)U = A_i U_{,i} \tag{3.61}$$

The same operator can be applied to the vector field $\mathbf{V}$. The following vector is obtained

$$(\mathbf{A} \cdot \mathrm{grad})V = (A_j \partial_j)(V_k \mathbf{i}_k) = A_j V_{k,j} \mathbf{i}_k \tag{3.62}$$

As a particular case, the total differential of a vector field $\mathbf{V} = \mathbf{V}(x_1, x_2, x_3)$ is obtained, under the form

$$d\mathbf{V} = (d\mathbf{r} \cdot \mathrm{grad})\mathbf{V} \tag{3.63}$$

Another special case is the directional derivative on the direction of unit vector $\mathbf{n}$, which has the form

$$\frac{\partial \mathbf{V}}{\partial n} = (\mathbf{n} \cdot \mathrm{grad})\mathbf{V} \tag{3.64}$$

In the case of a curve parameterized by the relationship $\mathbf{r} = \mathbf{r}(s)$, one uses $\mathbf{n} = \boldsymbol{\tau}$ and the result is

$$\frac{\partial \mathbf{V}}{\partial s} = \frac{d\mathbf{V}}{ds} = (\boldsymbol{\tau} \cdot \mathrm{grad})\mathbf{V} \tag{3.65}$$

3.2.1.3 First Order Differential Operators

Consider again the vector A, constant or variable. It may be used to define the vector differential operator $\mathbf{A} \times \nabla$:

$$\mathbf{A} \times \nabla \equiv \in_{jkl} A_j \partial_k \mathbf{i}_l \tag{3.66}$$

Applying this operator on a scalar field U, the result is

$$(\mathbf{A} \times \nabla)U = \in_{jkl} A_j U_{,k} \mathbf{i}_l \tag{3.67}$$

The divergence of a vector $\mathbf{V}$ is defined by using the differential operator $\nabla \cdot$, applied to the vector $\mathbf{V}$. The following equivalent notations are often used:

$$\mathrm{div}\mathbf{V} \equiv \nabla \cdot \mathbf{V} = \partial_i V_i = V_{i,i} \tag{3.68}$$

The *divergence* is a scalar quantity and is therefore invariable for a change in the system of coordinates.

The *curl* of a vector **V** is defined by using the differential operator $\nabla\times$, applied to the vector **V**:

$$\mathrm{curl}\mathbf{V} \equiv \nabla \times \mathbf{V} = \in_{jkl} \partial_j V_k \mathbf{i}_l = \in_{jkl} V_{k,j} \mathbf{i}_l \tag{3.69}$$

A vector field is called *irrotational* if

$$\mathrm{curl}\mathbf{V} = \nabla \times \mathbf{V} = 0 \tag{3.70}$$

It is noted that the gradient field:

$$\mathbf{V} = \mathrm{grad}\, U = \nabla U \tag{3.71}$$

is irrotational because $\mathrm{curl}\,\mathrm{grad}\, U = 0$. It follows that all fields of conservative vectors are at the same time irrotational. Moreover, these fields are the only ones that have this property. The property is valid for quasi-conservative fields because the condition (3.70) is equivalent to conditions (3.39). This is the condition that the integrand in the integral (3.53) represents an exact total differential.

A vector field is called *solenoidal* if

$$\mathrm{div}\mathbf{V} = \nabla \cdot \mathbf{V} = 0 \tag{3.72}$$

It is noted that a field of curls **V**, defined as:

$$\mathbf{V} = \mathrm{curl}\mathbf{W} = \nabla \times \mathbf{W} \tag{3.73}$$

is always solenoidal, because $\mathrm{div}\,\mathrm{curl}\, W = 0$. It can be shown that this is the only vector field which is solenoidal.

3.3 Elements of Exterior Differential Calculus

The main source further used for the exterior differential calculus is Mihaileanu (1972, p. 130). Consider the following linear differential form:

$$\omega = \lambda dx + \mu dy + v dz \tag{3.74}$$

where λ, μ, v are functions of x, y, z. An expression ω of the type (3.74) is called *Pfaff form* (or *one-form*). Consider two differentiation directions, specified by the letters d and δ, respectively:

$$\begin{gathered} \omega(d) = \sum_i \lambda_i dx^i, \qquad \omega(\delta) = \sum_i \lambda_i \delta x^i \\ (\lambda_i = \lambda, \mu, v), \quad (x^i = x, y, z) \end{gathered} \tag{3.75}$$

The alternative differentials are:

$$\begin{aligned} \delta\omega(d) &= \sum_i \left(\frac{\partial\lambda_i}{\partial x}\delta x + \frac{\partial\lambda_i}{\partial y}\delta y + \frac{\partial\lambda_i}{\partial z}\delta z \right) dx^i + \sum_i \lambda_i \delta dx^i \\ d\omega(\delta) &= \sum_i \left(\frac{\partial\lambda_i}{\partial x}dx + \frac{\partial\lambda_i}{\partial y}dy + \frac{\partial\lambda_i}{\partial z}dz \right) \delta x^i + \sum_i \lambda_i d\delta x^i \end{aligned} \tag{3.76a, b}$$

If the operators d, δ are permutable, by applying them to the independent variables one finds:

$$\delta dx = d\delta x, \quad \delta dy = d\delta y, \quad \delta dz = d\delta z \tag{3.77}$$

By subtracting relations (3.76), it is obtained:

$$\begin{aligned} \delta\omega(d) - d\omega(\delta) &= \sum_{i,j} \frac{\partial\lambda_i}{\partial y^j}\left(dx^i \delta y^j - dy^j dx^i\right) \\ &= \frac{1}{2}\sum_{i,j} \left(\frac{\partial\lambda_i}{\partial y^j} - \frac{\partial\mu_j}{\partial x^i} \right) \left(dx^i \delta y^j - dy^j \delta x^i\right) \end{aligned} \tag{3.78}$$

Denote:

$$\begin{aligned} [\omega_1, \omega_2] &= \begin{vmatrix} \sum\limits_i \lambda_{1i} dx^i & \sum\limits_i \lambda_{2i} dx^i \\ \sum\limits_i \lambda_{1i} \delta x^i & \sum\limits_i \lambda_{2i} \delta x^i \end{vmatrix} \\ d\omega &= \delta\omega(d) - d\omega(\delta) \end{aligned} \tag{3.79, 80}$$

From (3.78) and (3.80) it is obtained:

$$d\omega = \frac{1}{2}\sum_{i,j} \left(\frac{\partial\lambda_i}{\partial y^j} - \frac{\partial\mu_j}{\partial x^i} \right) \left[dx^i, dy^i\right] \tag{3.81}$$

Relation (3.81) allows the introduction of two important notions. First, $[\omega_1, \omega_2]$ is called the *exterior product* of the forms $\omega_k = \sum\limits_i \lambda_{ki} dx^i$. Secondly, $d\omega$ is the *exterior differential* of the form ω. From (3.81) it is obvious that if ω is an exact differential, then $d\omega = 0$. The following relationship is true:

$$[\omega_1, \omega_2] = -[\omega_2, \omega_1] \tag{3.82}$$

The exterior differential and the exterior product have the following properties:

$$\begin{aligned}
&[\lambda\omega_1, \omega_2] = \lambda[\omega_1, \omega_2] \\
&[\omega_1, (\omega_2 + \omega_3)] = [\omega_1, \omega_2] + [\omega_1, \omega_3] \\
&d(\lambda\omega) = \lambda d\omega + [d\lambda, \omega] \\
&d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2 \\
&d(d\omega) = 0
\end{aligned} \tag{3.83a–e}$$

The exterior differential calculus was introduced by Cartan in 1926. The important theorem (3.83e) has been demonstrated by Henry Poincaré.

3.3.1 Equations of Space Structure

Assume that $\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3$ is a three-orthogonal reference frame. By definition, for this reference frame the following relations apply:

$$\mathbf{I}_1^2 = \mathbf{I}_2^2 = \mathbf{I}_3^2 = 1 \qquad \mathbf{I}_1 \cdot \mathbf{I}_2 = \mathbf{I}_2 \cdot \mathbf{I}_3 = \mathbf{I}_3 \cdot \mathbf{I}_1 = 0 \tag{3.84}$$

Consider an arbitrary point $\mathrm{M}(x, y, z)$. Consider an infinitesimal movement of the reference frame. The movement $d\mathbf{I}$ is accompanied by a variation $d\mathbf{M}$ of the position vector of the point M. The following relationships are true:

$$\begin{aligned}
d\mathbf{M} &= \omega^1\mathbf{I}_1 + \omega^2\mathbf{I}_2 + \omega^3\mathbf{I}_3 \\
d\mathbf{I} &= \omega_i^1\mathbf{I}_1 + \omega_i^2\mathbf{I}_2 + \omega_i^3\mathbf{I}_3
\end{aligned} \tag{3.85, 86}$$

The following relationship is obtained by differentiation of (3.84):

$$\mathbf{I}_1 d\mathbf{I}_2 + \mathbf{I}_2 d\mathbf{I}_1 = 0 \tag{3.87}$$

Taking into account (3.86) all terms will become equal to zero, except $\mathbf{I}_1^2$ and $\mathbf{I}_2^2$, which involve the coefficients ω_2^1 and ω_1^2. The same occurs for the homologous products. Therefore

$$\omega_j^i + \omega_i^j = 0 \tag{3.88}$$

In particular, $\omega_i^i = 0$. Of the nine coefficients $\omega_i^j (i,j = 1, 2, 3)$, only three independent coefficients remain, for example

$$\omega_2^1 = u^3; \quad \omega_3^2 = u^1; \quad \omega_1^3 = u^2 \tag{3.89}$$

which cause the rotation of the reference frame. To these one can add the three parameters $\omega^1, \omega^2, \omega^3$ that determine the translation of the reference frame.

The exterior differential is applied to the relationship (3.85). By using the foregoing properties, it is obtained:

$$\begin{aligned}
0 = d(d\mathbf{M}) = d\left(\sum_k \omega^k I_k\right) &= \sum_k d(\omega^k I_k) = \sum_k \left\{I_k d\omega^k + \left[dI_k, \omega^k\right]\right\} \\
&= \sum_k \left\{I_k d\omega^k + \sum_h \left[\omega_k^h I_h, \omega^k\right]\right\} = \sum_k \left\{I_k d\omega^k + \sum_h [\omega_k^h, \omega^k] I_h\right\} \\
&= \sum_k I_k \left\{d\omega^k + \sum_h \left[\omega_h^k, \omega^h\right]\right\}
\end{aligned} \tag{3.90}$$

Since the vectors $\mathbf{I}_k$ are independent, one obtains:

$$d\omega^k = \sum_h \left[\omega^h, \omega_h^k\right] \qquad (k = 1, 2, 3) \tag{3.91}$$

By differentiation of formula (3.86), it follows:

$$\begin{aligned}
0 = d(d\mathbf{I}_i) = d\left(\sum_k \omega_i^k I_k\right) &= \sum_k d\left(\omega_i^k I_k\right) \\
&= \sum_k \left\{I_k d\omega_i^k + \left[dI_k, \omega_i^k\right]\right\} = \sum_k \left\{I_k d\omega_i^k + \sum_h \left[\omega_k^h I_h, \omega_i^k\right]\right\} \\
&= \sum_k \left\{I_k d\omega_i^k + \sum_h \left[\omega_k^h, \omega_i^k\right] I_h\right\} = \sum_k I_k \left\{d\omega_i^k + \sum_h \left[\omega_h^k, \omega_i^h\right]\right\}
\end{aligned} \tag{3.92}$$

Therefore

$$d\omega_i^k = \sum_h \left[\omega_i^h, \omega_h^k\right] \qquad (i, h = 1, 2, 3.) \tag{3.93}$$

Equations (3.91) and (3.93) are the so-called *Cartan equations of the space structure*.

3.3.2 Applications to the Theory of Surfaces

The relations (3.86) become:

$$d\mathbf{I}_1 = -u^3\mathbf{I}_2 + u^2\mathbf{I}_3 \quad d\mathbf{I}_2 = u^3\mathbf{I}_1 - u^1\mathbf{I}_3 \quad d\mathbf{I}_3 = -u^2\mathbf{I}_1 + u^1\mathbf{I}_2 \tag{3.94}$$

Use the Cartan equations of the space structure. Thus, relations (3.91) become

$$\begin{aligned} &d\omega^1 = \left[\omega^2, u^3\right] - \left[\omega^3, u^2\right], \quad d\omega^2 = \left[\omega^3, u^1\right] - \left[\omega^1, u^3\right], \\ &d\omega^3 = \left[\omega^1, u^2\right] - \left[\omega^3, u^1\right] \end{aligned} \tag{3.95}$$

and from the relationships (3.93) one finds:

$$du^1 = -\left[u^2, u^3\right], \quad du^2 = -\left[u^3, u^1\right], \quad du^3 = -\left[u^1, u^2\right] \tag{3.96}$$

The notion of exterior product will be used in subsequent chapters and it will obtain a richer physical meaning.

References

Beju, I., Soos, E., Teodorescu, P.P.: Tehnici de calcul vectorial cu aplicatii, Editura Tehnica, Bucuresti (1976)

Mihaileanu, N.: Geometrie analitica, proiectiva si diferentiala. Complemente, Ed. Didactica si Pedagogica, Bucuresti (1972)

Chapter 4
Elements of Riemann Geometry

There are many bibliographic sources that can be used for a good introduction to the Riemann spaces and Riemann geometry. Here the guide is the work of Beju et al. (Tehnici de calcul vectorial cu aplicatii, Ed. Tehnica, Bucharest, 1976), which has the advantage of a more pronounced intuitive character. Some notions will be treated again on a more abstract basis in the second section of this chapter, and in subsequent chapters.

4.1 Standard Notions of Non-Euclidean Geometry

4.1.1 Riemann Spaces

Consider a coordinate system $\{x^i\}$ in a n-dimensional space. Define the *fundamental form*, which is a positive definite quadratic form, whose coefficients are differentiable functions of the coordinates $\{x^i\}$. It is said that the n-dimensional space is a metric space. This is justified by the fact that introducing the fundamental form is equivalent to defining the element of arc length:

$$ds^2 = g_{ij} dx^i dx^j \tag{4.1}$$

Such a metric space is called n-dimensional Riemann space (denoted V_n). Considered another coordinate system, $\{x'^i\}$. Assume that there is a non-singular coordinate transformation $\{x^i\} \to \{x'^i\}$:

$$x'^j = x'^j(x^1, x^2, \ldots, x^n) \quad (j = 1, \ldots, n) \tag{4.2}$$

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_4

which yields to the next form of the arc length element, as function of the new coordinates:

$$ds^2 = (dx'^1)^2 + (dx'^2)^2 + \cdots + (dx'^n)^2 \tag{4.3}$$

Then, it is said that the *Riemann space* V_n is an *Euclidean space* E_n. Euclidean spaces are frequently used in applications.

4.1.2 Curvilinear Coordinates with Arbitrary Bases

Assume an arbitrary point P, which will be specified by its vector of position **r**. In general, the position vector is a function of coordinates, which are denoted (q^1, q^2, q^3) in the three-dimensional case. Depending on the coordinates, the position vector is written as follows:

$$\mathbf{r}(q^1, q^2, q^3) = x_j(q^1, q^2, q^3)\mathbf{i}_j \tag{4.4}$$

In the case that the moving point (q^1, q^2, q^3) generates a domain D of R^3, the point P generates a domain denoted V. In every point of the domain V one can build three coordinate curves: the first curve is given by $q^2 = \text{const}$ and $q^3 = \text{const}$, the second curve is given by $q^3 = \text{const}$ and $q^1 = \text{const}$ and the third curve is given by $q^1 = \text{const}$ and $q^2 = \text{const}$. The coordinates on these curves are called *curvilinear coordinates*. Between the orthogonal Cartesian coordinates (x_1, x_2, x_3) and the curvilinear coordinates, there is the following relation

$$x_j = x_j(q^1, q^2, q^3) \quad (j = 1, 2, 3) \tag{4.5}$$

In the following it will be assumed that the functions occurring in (4.5) are differentiable. In the case that

$$\det\left|\frac{\partial(x_1, x_2, x_3)}{\partial(q^1, q^2, q^3)}\right| \neq 0 \tag{4.6}$$

the transformation (4.5) from the curvilinear coordinates to the orthogonal Cartesian coordinates is invertible. In ordinary cases it is assumed that this transformation is biunivocal. This is equivalent to saying that to a point $\mathrm{P}(\mathrm{x}_1, x_2, x_3)$ corresponds a single system of curvilinear coordinates (q^1, q^2, q^3), and vice versa.

The differential of the position vector is given by:

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial q^i} dq^i \tag{4.7}$$

Note that the vectors $\mathbf{e}^i (i = 1, 2, 3)$, defined as follows:

$$\mathbf{e}^i \equiv \frac{\partial \mathbf{r}}{\partial q^i} \quad (i = 1, 2, 3) \tag{4.8}$$

are tangent to the coordinate curves. These vectors constitute a local basis because

$$(e^1, e^2, e^3) = \det\left|\frac{\partial(x_1, x_2, x_3)}{\partial(q^1, q^2, q^3)}\right| \neq 0 \tag{4.9}$$

The traditional notation, already used in the previous chapter, is used in the following to define the scalar product between pairs of vectors of the local basis:

$$g_{ij} = g_{ji} = \frac{\partial \mathbf{r}}{\partial q^i} \cdot \frac{\partial \mathbf{r}}{\partial q^j} = \mathbf{e}_i \cdot \mathbf{e}_j \quad (i, j = 1, 2, 3) \tag{4.10}$$

4.1.3 Curvilinear Coordinates with Orthogonal Bases

The maximum number of different quantities g_{ij} $(i, j = 1, 2, 3)$ is nine. An important particular case is that of a curvilinear orthogonal coordinate system, for which:

$$\begin{aligned} &g_{ij} = 0, \quad i \neq j \\ &g_{11} = \left(\frac{\partial \mathbf{r}}{\partial q^1}\right)^2, \quad g_{22} = \left(\frac{\partial \mathbf{r}}{\partial q^2}\right)^2, \quad g_{33} = \left(\frac{\partial \mathbf{r}}{\partial q^3}\right)^2 \end{aligned} \tag{4.11}$$

Two cases of curvilinear orthogonal coordinates are often encountered in practice, i.e. the cylindrical coordinates and the spherical coordinates, which are briefly defined below.

A cylindrical coordinate system (r, θ, z), with $r \geq 0,\ 0 \leq \theta \leq 2\pi$, is linked to the orthogonal Cartesian coordinates of the point P through the relationships

$$x_1 = r \cos\theta, \quad x_2 = r \sin\theta, \quad x_3 = z \tag{4.12}$$

A spherical coordinate system (R, φ, θ), with $R \geq 0,\ 0 \leq \varphi \leq \pi,\ 0 \leq \theta \leq 2\pi$, is linked to the orthogonal Cartesian coordinates through the relations

$$x_1 = R \sin\varphi \cos\theta, \quad x_2 = R \sin\varphi \sin\theta, \quad x_3 = R \cos\varphi \tag{4.13}$$

4.1.4 Element of Volume

By using the vectors $\mathbf{e}_1 dq^1, \mathbf{e}_2 dq^2, \mathbf{e}_3 dq^3$ one can build a curvilinear parallelepiped. The element of volume of this parallelepiped is given by

$$d\tau = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) dq^1 dq^2, dq^3 \tag{4.14}$$

If the vectors $\mathbf{e}^i (i = 1, 2, 3)$ constitute a positive basis, one may show that the volume element is given by

$$d\tau = g^{1/2} dq^1 dq^2 dq^3 \tag{4.15}$$

In the particular case of orthogonal curvilinear coordinates, there are major simplifications of calculation. One of these simplifications is:

$$g = g_{11} g_{22} g_{33} \tag{4.16}$$

Also, it is easily seen that the expression of the volume element in cylindrical coordinates is

$$d\tau = r dr d\theta dz \tag{4.17}$$

and in spherical coordinates it is

$$d\tau = R^2 \sin \varphi dR d\varphi d\theta \tag{4.18}$$

4.1.5 Element of Arc Length

Consider the case of a point P moving on a curve C, that will be parameterized using a parameter $t \in [t_0, t_1]$ and three mappings $t \to q^i(t) (i = 1, 2, 3)$. The element of arc length on the curve C, $ds \equiv |d\mathbf{r}|$, is defined as follows

$$ds^2 = d\mathbf{r}^2 = \frac{\partial \mathbf{r}}{\partial q^i} dq^i \cdot \frac{\partial \mathbf{r}}{\partial q^i} dq^i \tag{4.19}$$

By using the notation (4.10), another very useful way to writing the square of the arc element is obtained:

$$ds^2 = g_{ij} dq^i dq^j \tag{4.20}$$

Relation (4.20) defines the metric on the Euclidean space.

In the case of orthogonal curvilinear coordinates, the expressions are significantly simplified. A few examples useful in applications are presented in the following. Thus, by direct calculation it can be shown that in cylindrical coordinates the arc element is given by the following expression

$$ds^2 = dr^2 + r^2 d\theta^2 + dz^2 \tag{4.21}$$

In spherical coordinates, the expression is:

$$ds^2 = dR^2 + R^2 d\varphi^2 + R^2 \sin^2 \varphi d\theta^2 \tag{4.22}$$

4.1.6 Tensors

The curvilinear coordinates and the transformations between different coordinate systems can be used to systematically classify the physical quantities.

4.1.6.1 Transformations of Curvilinear Coordinates

Consider two curvilinear coordinate systems: q^i $(i = 1, 2, 3)$ and $q'^i(i = 1, 2, 3)$, respectively. The bases of these systems are $\{\mathbf{e}_i\}$ and $\{\mathbf{e}'_i\}$, respectively. The transformation from the first coordinate system to the second one is done by using the mappings $(q^1, q^2, q^3) \to q'^j(q^1, q^2, q^3)$ $(j = 1, 2, 3)$, assumed as being differentiable. According to (4.8), the basis $\{\mathbf{e}'_i\}$ of the second coordinate system is given by

$$\mathbf{e}'_j = \frac{\partial \mathbf{r}}{\partial q'^j} \tag{4.23}$$

Using these relationships for the basis of the first coordinate system leads to the following relations

$$\mathbf{e}'_j = \mathbf{e}_i \frac{\partial q^i}{\partial q'^j} \quad (j = 1, 2, 3) \tag{4.24}$$

Similarly, the following relationships are obtained

$$\mathbf{e}_i = \mathbf{e}'_j \frac{\partial q'^j}{\partial q^i} \quad (i = 1, 2, 3) \tag{4.25}$$

Multiplying the expressions (4.24) and (4.25) among themselves and taking into account the properties of the scalar product between the vectors of the bases, it is found that:

$$\frac{\partial q^i}{\partial q'^j}\frac{\partial q'^j}{\partial q^k}=\delta^i_k,\quad \frac{\partial q'^i}{\partial q^j}\frac{\partial q^j}{\partial q'^k}=\delta^i_k \tag{4.26}$$

4.1.6.2 Tensors of Zero Order

It is said that a function $U(q^1,q^2,q^3)$ is a zero-order tensor (or a scalar) if it is invariant under a change of curvilinear coordinates. This means that through a change of curvilinear coordinates, the function $U(q^1,q^2,q^3)$ is transformed into the function $U'(q'^1,q'^2,q'^3)$, by obeying the condition of invariance.

Relative Tensors of Zero Order

Denote by J the Jacobian of the curvilinear coordinate transformation $(q^1,q^2,q^3)\rightarrow q'^j(q^1,q^2,q^3)$ $(j=1,2,3)$:

$$J\equiv \det\left|\frac{\partial(q'^1,q'^2,q'^3)}{\partial(q^1,q^2,q^3)}\right| \tag{4.27}$$

The coefficients of the fundamental form are introduced in the second coordinate system, under the form of the scalar products

$$g'_{ij}=\mathbf{e}'_i\cdot\mathbf{e}'_j\quad (i,j=1,2,3) \tag{4.28}$$

Denote $g'\equiv\det|g'_{ij}|$. Also, the following relationship takes place:

$$(g')^{1/2}=\frac{1}{J}(g)^{1/2} \tag{4.29}$$

It is said that a function U is a *relative scalar* of *weighting* P if, after a change of curvilinear coordinate, it is transformed according to the following relationship

$$U'=J^{-P}U \tag{4.30}$$

In case $P=0$, it is an *absolute scalar*, if $P=1$ it is a *scalar density* and if $P=-1$ it is a *scalar capacity*.

4.1.6.3 First Order Tensors (Vectors)

Contravariant Components

In the following, formulas are obtained for the transformation of the contravariant components of a vector in the basis $\{\mathbf{e}_i\}$ into contravariant components in the basis $\{\mathbf{e}'_i\}$, and vice versa. Note that a given vector can be represented in the two bases in the following forms

$$\mathbf{V} = V'^j \mathbf{e}'_j = V^i \mathbf{e}_i = V'^j \frac{\partial q^i}{\partial q'^j} \mathbf{e}_i = V^i \frac{\partial q'^j}{\partial q^i} \mathbf{e}'_j \tag{4.31}$$

From the previous equations, the next expressions are obtained:

$$V'^j = V^i \frac{\partial q'^j}{\partial q^i} (j = 1, 2, 3) \quad V^i = V'^j \frac{\partial q^i}{\partial q'^j} (i = 1, 2, 3) \tag{4.32, 33}$$

The sets of threes functions which obey the formulas (4.32) or (4.33) in a transformation of curvilinear coordinates are the contravariant components of a *first-order tensor* (or the contravariant components of a vector).

Covariant Components

Next, formulas are obtained for the transformation of the covariant components of a vector in the basis $\{\mathbf{e}_i\}$ into the covariant components in the base $\{\mathbf{e}'_i\}$, and vice versa. To do this, the definition of the mixed product of three vectors is used and Eq. (4.24) is taken into account. It is obtained

$$(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \det \left| \frac{\partial(q^1, q^2, q^3)}{\partial(q'^1, q'^2, q'^3)} \right| = \frac{1}{J} (g)^{1/2} \tag{4.34}$$

The vectors of the reciprocal basis of a curvilinear coordinate system can be expressed by using Eq. (3.11). By using the normalized minors of the elements of the functional determinant given by (4.27), the following two sets of relationships are obtained:

$$\mathbf{e}'^j = \mathbf{e}^i \frac{\partial q'^j}{\partial q^i} \quad (j = 1, 2, 3); \quad \mathbf{e}^i = \mathbf{e}'^j \frac{\partial q^i}{\partial q'^j} \quad (i = 1, 2, 3) \tag{4.35, 36}$$

A vector $\mathbf{V}$ can be represented by means of the covariant components:

$$\mathbf{V} = V'_j \mathbf{e}'^j = V_i \mathbf{e}^i = V'_j \frac{\partial q'^j}{\partial q^i} \mathbf{e}_i = V_i \frac{\partial q^i}{\partial q'^j} \mathbf{e}'^j \tag{4.37}$$

This yields the following two series of formulas:

$$V'_j = V_i \frac{\partial q^i}{\partial q'^j} \quad (j = 1,2,3); \quad V_i = V'_j \frac{\partial q'^j}{\partial q^i} \quad (i = 1,2,3) \qquad (4.38, 39)$$

The sets of threes functions which obey the formulas (4.38) or (4.39) in a transformation of curvilinear coordinates are the covariant components of a *first-order tensor* (or covariant components of a vector).

Relative Tensors of the First Order

The sets of three functions which obey the following formulas in a transformation of curvilinear coordinates:

$$V'_j = J^{-P} V_i \frac{\partial q^i}{\partial q'^j} \quad (j = 1,2,3) \qquad (4.40)$$

are the contravariant components of a relative tensor of first-order (or, in other words, the contravariant components of a *relative vector*). The case $P = 0$ corresponds to the contravariant components of an *absolute vector*, the case $P = 1$ corresponds to a *vector density*, and the case $P = -1$ corresponds to a *vector capacity*. Similarly, one may define the covariant components of a relative vector.

The transformations of coordinates treated above are linear and homogeneous. This can be demonstrated by using the transitivity property of the Jacobian of the transformation:

$$\det\left|\frac{\partial(q''^1, q''^2, q''^3)}{\partial(q^1, q^2, q^3)}\right| = \det\left|\frac{\partial(q''^1, q''^2, q''^3)}{\partial(q'^1, q'^2, q'^3)}\right| \det\left|\frac{\partial(q'^1, q'^2, q'^3)}{\partial(q^1, q^2, q^3)}\right| \qquad (4.41)$$

Therefore, if all the components of a relative vector are zero in a reference frame, they also become equal to zero in any other reference frame. An important consequence is: if an equation containing relative vectors is true in a curvilinear coordinate system, it is true in any other curvilinear coordinate system. Also, the relative vectors in the left-hand side and the right-hand side of an equation, respectively, must have the same weighting.

4.1.7 Differential Operators in Curvilinear Coordinates

If the condition (4.6) is met, from the relations (4.5) one may deduce the existence of some relationships of the form

$$q^i = q^i(x_1, x_2, x_3) \quad (i = 1, 2, 3) \tag{4.42}$$

4.1.7.1 First Order Differential Parameters

Denote by $\{\mathbf{e}_i\}$ and $\{\mathbf{e}^i\}$ the vectors of the basis and reciprocal basis, respectively. By defining some surfaces of coordinates $q^i = \text{const} \quad (\text{i} = 1, 2, 3)$, one arrives to the following form of the gradient of the curvilinear coordinates given by the relationships (4.42):

$$\text{grad } q^i = \mathbf{e}^i \quad (i = 1, 2, 3) \tag{4.43}$$

The modules of these gradients of coordinates are called *first order differential parameters* and are given by the following relationships containing coefficients of the fundamental form associated with the basis $\{\mathbf{e}^i\}$:

$$\begin{gathered} h^1 \equiv |\text{grad } q^1| = (g^{11})^{\frac{1}{2}}, \qquad h^2 \equiv |\text{grad } q^2| = (g^{22})^{\frac{1}{2}}, \\ h^3 \equiv |\text{grad } q^3| = (g^{33})^{\frac{1}{2}} \end{gathered} \tag{4.44}$$

The unit vectors of the reciprocal basis $\{\mathbf{e}^i\}$ are compactly expressed by using the first order differential parameters, as follows

$$\mathbf{i}^1 = \frac{1}{h^1}\mathbf{e}^1, \quad \mathbf{i}^2 = \frac{1}{h^2}\mathbf{e}^2, \quad \mathbf{i}^3 = \frac{1}{h^3}\mathbf{e}^3 \tag{4.45}$$

4.1.7.2 Lamé Coefficients

The coefficients of the fundamental form associated with the basis $\{\mathbf{e}_i\}$ can be used to define the *Lamé coefficients*, whose squares are

$$H_1^2 \equiv g_{11}, \quad H_2^2 \equiv g_{22}, \quad H_3^2 \equiv g_{33} \tag{4.46}$$

The Lamé coefficients allow to writing the unit vectors of the basis $\{\mathbf{e}_i\}$ under the form

$$\mathbf{i}_1 = \frac{1}{H_1}\mathbf{e}_1, \quad \mathbf{i}_2 = \frac{1}{H_2}\mathbf{e}_2, \quad \mathbf{i}_3 = \frac{1}{H_3}\mathbf{e}_3 \tag{4.47}$$

The physical components of a vector $\mathbf{V}$ correspond to the three sides of the parallelepiped built on the three vectors of the basis $\{\mathbf{e}_i\}$, having $\mathbf{V}$ as a diagonal. The Lamé coefficients allow to writing the expression of these physical components in the compact form (H_1V^1, H_2V^2, H_3V^3). For this, see the relationship (3.9).

Similarly, the orthogonal projections of the vector **V** on the three vectors of the basis $\{\mathbf{e}_i\}$ are given by $(V_1/H_1, V_2/H_2, V_3/H_3)$.

4.1.7.3 Conversion Formulas

Using the relations (3.11), (3.24), (4.47) and (4.45) one obtains:

$$h^1 = \frac{1}{(g)^{1/2}} H_2 H_3, \quad h^2 = \frac{1}{(g)^{1/2}} H_3 H_1, \quad h^3 = \frac{1}{(g)^{1/2}} H_1 H_2 \tag{4.48}$$

$$H_1^2 = (g)^{1/2} \frac{h^2 h^3}{h^1}, \quad H_2^2 = (g)^{1/2} \frac{h^3 h^1}{h^2}, \quad H_3^2 = (g)^{1/2} \frac{h^1 h^2}{h^3} \tag{4.49}$$

These relationships represent conversion formulas between the first order differential parameters and the Lamé coefficients, and vice versa. The differential parameters and the Lamé coefficients are used for convenient switching from a Cartesian orthogonal reference frame (x_1, x_2, x_3) to a system of curvilinear coordinates (q_1, q_2, q_3).

4.1.7.4 Element of Arc Length

The arc length element is given by the general formula (4.20). Its components along the coordinate lines are

$$ds^1 = H_1 dq^1, \quad ds^2 = H_2 dq^2, \quad ds^3 = H_3 dq^3, \tag{4.50}$$

4.1.7.5 Gradient

The gradient of a scalar field $U(q^1, q^2, q^3)$ can be expressed as

$$\mathrm{grad}U = \frac{\partial U}{\partial q^i} \mathrm{grad}\, q^i = \frac{\partial U}{\partial q^i} \mathbf{e}^i \tag{4.51}$$

Therefore, $\partial U / \partial q^i$ are the covariant components of the gradient in the system of curvilinear coordinates. Consequently, in a curvilinear coordinate system having the basis $\{\mathbf{e}^i\}$ and the unit vectors $\mathbf{i}^j (j = 1, 2, 3)$, the gradient operator is given by

$$\mathrm{grad} = h^1 \mathbf{i}^1 \frac{\partial}{\partial q^1} + h^2 \mathbf{i}^2 \frac{\partial}{\partial q^2} + h^3 \mathbf{i}^3 \frac{\partial}{\partial q^3} \tag{4.52}$$

4.1.7.6 Divergence and Curl

The relations (3.68) and (3.69) define the divergence and the curl of a vector, respectively. In curvilinear coordinates, the divergence and the curl can be expressed in the following form

$$\mathrm{div}V = \frac{1}{(g)^{1/2}} \frac{\partial}{\partial q^i} [(g)^{1/2} V^i] \tag{4.53}$$

$$\mathrm{curl}\mathbf{V} = \frac{1}{(g)^{1/2}} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial q^1} & \frac{\partial}{\partial q^2} & \frac{\partial}{\partial q^3} \\ V_1 & V_2 & V_3 \end{vmatrix} = \frac{1}{(g)^{1/2}} \in^{ijk} \frac{\partial V_k}{\partial q^j} \mathbf{e}_i \tag{4.54}$$

4.1.7.7 Orthogonal Curvilinear Coordinates

In practice, systems of orthogonal curvilinear coordinates are often used. In this important case, the reciprocal basis $\{\mathbf{e}_i\}$ is orthogonal, while its vectors have the same directions as the vectors of the basis $\{\mathbf{e}^i\}$. Also, $g_{ij} = g^{ij} = 0$ for $i \neq j$. Consequently:

$$g = (H_1 H_2 H_3)^2 = \frac{1}{(h^1 h^2 h^3)^2} \tag{4.55}$$

In general $g \neq 1$, because the basis $\{\mathbf{e}_i\}$, having $\mathbf{i}^1 = \mathbf{i}_1$, $\mathbf{i}^2 = \mathbf{i}_2$, $\mathbf{i}^3 = \mathbf{i}_3$, is not orthonormal. Relations (4.49) lead to

$$H_1 = \frac{1}{h^1}, \quad H_2 = \frac{1}{h^2}, \quad H_3 = \frac{1}{h^3} \tag{4.56}$$

Detailed calculations for quantities often encountered in practice are presented next. Two particular cases will be considered, namely cylindrical coordinates and spherical coordinates, respectively.

Element of Arc Length

The arc length element is given by the expression

$$\begin{aligned} ds^2 &= (ds^1)^2 + (ds^2)^2 + (ds^3)^2 \\ &= H_1^2 (dq^1)^2 + H_2^2 (dq^2)^2 + H_3^2 (dq^3)^2 \end{aligned} \tag{4.57}$$

In cylindrical coordinates, the following simple relationships take place:

$$H_1 = 1, \quad H_2 = r \quad , H_3 = 1, \quad g = r^2 \tag{4.58}$$

For the line of coordinates $q^2 = \theta$, the expression of the component of the arc length element can be written as

$$ds^\theta = rd\theta, \tag{4.59}$$

In spherical coordinates, the following relations take place

$$H_1 = 1, \quad H_2 = R, \quad H_3 = R\sin\varphi, \quad g = R^4\sin^2\varphi \tag{4.60}$$

For the coordinates $q^2 = \varphi$, $q^3 = \theta$, the following expressions of the components of the arc length element are obtained

$$ds^\varphi = Rd\varphi, \quad ds^\theta = R\sin\varphi d\theta, \tag{4.61}$$

Element of Volume

The element of volume, generally given relation (4.15), takes the very simple form

$$d\tau = H_1H_2H_3dq^1dq^2dq^3 \tag{4.62}$$

Gradient

The gradient operator, which has the general form (4.52), becomes in this particular case:

$$\mathrm{grad} = \frac{1}{H_1}\mathbf{i}_1\frac{\partial}{\partial q^1} + \frac{1}{H_2}\mathbf{i}_2\frac{\partial}{\partial q^2} + \frac{1}{H_3}\mathbf{i}_3\frac{\partial}{\partial q^3} \tag{4.63}$$

In the case of the cylindrical coordinates, the next relationship occurs:

$$\frac{\partial}{\partial s^\theta} = \frac{1}{r}\frac{\partial}{\partial\theta} \tag{4.64}$$

and the gradient operator is reduced to:

$$\mathrm{grad} = \mathbf{i}_r\frac{\partial}{\partial r} + \frac{1}{r}\mathbf{i}_\theta\frac{\partial}{\partial\theta} + \mathbf{i}_z\frac{\partial}{\partial z} \tag{4.65}$$

If spherical coordinates are considered, the following relations occur:

$$\frac{\partial}{\partial s^{\varphi}}=\frac{1}{R}\frac{\partial}{\partial\varphi}, \quad \frac{\partial}{\partial s^{\theta}}=\frac{1}{R\sin\varphi}\frac{\partial}{\partial\theta} \tag{4.66}$$

and the gradient operator has the form

$$\mathrm{grad}=\mathbf{i}_R\frac{\partial}{\partial R}+\frac{1}{R}\mathbf{i}_{\varphi}\frac{\partial}{\partial\varphi}+\frac{1}{R\sin\varphi}\mathbf{i}_{\theta}\frac{\partial}{\partial\theta} \tag{4.67}$$

4.1.7.8 Divergence

In curvilinear orthogonal coordinates, the divergence has the following form:

$$\begin{aligned}\mathrm{div}\mathbf{V}=\frac{1}{H_1H_2H_3}\bigg[&\frac{\partial}{\partial q^1}(H_2H_3V^1)+\frac{\partial}{\partial q^2}(H_3H_1V^2)\\&+\frac{\partial}{\partial q^3}(H_1H_2V^3)\bigg]\end{aligned} \tag{4.68}$$

In the particular cases of cylindrical coordinates and spherical coordinates, respectively, the divergence is expressed as:

$$\begin{aligned}\mathrm{div}\mathbf{V}&=\frac{1}{r}\frac{\partial}{\partial r}(rV^r)+\frac{1}{r}\frac{\partial V^{\theta}}{\partial\theta}+\frac{\partial V^z}{\partial z}\\ \mathrm{div}\mathbf{V}&=\frac{1}{R^2}\frac{\partial}{\partial R}(R^2V^R)+\frac{1}{R\sin\varphi}\frac{\partial}{\partial\varphi}(\sin\varphi\, V^{\varphi})+\frac{1}{R\sin\varphi}\frac{\partial V^{\theta}}{\partial\theta}\end{aligned} \tag{4.69, 70}$$

4.1.7.9 Curl

In curvilinear orthogonal coordinates, the curl operator has the expression

$$\begin{aligned}\mathrm{curl}\mathbf{V}=&\frac{1}{H_2H_3}\left[\frac{\partial}{\partial q^2}(H_3V^3)-\frac{\partial}{\partial q^3}(H_2V^2)\right]\mathbf{i}_1\\&+\frac{1}{H_3H_1}\left[\frac{\partial}{\partial q^3}(H_1V^1)-\frac{\partial}{\partial q^1}(H_3V^3)\right]\mathbf{i}_2\\&+\frac{1}{H_1H_2}\left[\frac{\partial}{\partial q^1}(H_2V^2)-\frac{\partial}{\partial q^2}(H_1V^1)\right]\mathbf{i}_3\end{aligned} \tag{4.71}$$

In the particular cases of cylindrical coordinates and spherical coordinates, respectively, the curl is given by the following expressions:

$$\begin{aligned}
\operatorname{curl}\mathbf{V} &= \left(\frac{1}{r}\frac{\partial V^r}{\partial \theta} - \frac{\partial V^\theta}{\partial z}\right)\mathbf{i}_r + \left(\frac{\partial V^r}{\partial z} - \frac{\partial V^z}{\partial r}\right)\mathbf{i}_\theta \\
&\quad + \frac{1}{r}\left[\frac{\partial}{\partial r}(rV^\theta) - \frac{\partial V^r}{\partial \theta}\right]\mathbf{i}_z \\
\operatorname{curl}\mathbf{V} &= \frac{1}{R\,\sin\varphi}\left[\frac{\partial}{\partial \varphi}(\sin\varphi\, V^\theta) - \frac{\partial V^\varphi}{\partial \theta}\right]\mathbf{i}_R \\
&\quad + \frac{1}{R}\left[\frac{1}{\sin\varphi}\frac{\partial V^R}{\partial \theta} - \frac{\partial}{\partial R}(RV^\theta)\right]\mathbf{i}_\varphi + \frac{1}{R}\left[\frac{\partial}{\partial R}(RV^\varphi) - \frac{\partial V^R}{\partial \varphi}\right]\mathbf{i}_\theta
\end{aligned} \tag{4.72, 73}$$

4.1.7.10 Laplace Operator

The Laplace operator in curvilinear orthogonal coordinates can be expressed as:

$$\begin{aligned}
\Delta = \frac{1}{H_1H_2H_3}&\left[\frac{\partial}{\partial q^1}\left(\frac{H_2H_3}{H_1}\frac{\partial}{\partial q^1}\right) + \frac{\partial}{\partial q^2}\left(\frac{H_3H_1}{H_2}\frac{\partial}{\partial q^2}\right)\right. \\
&\left. + \frac{\partial}{\partial q^3}\left(\frac{H_1H_2}{H_3}\frac{\partial}{\partial q^3}\right)\right]
\end{aligned} \tag{4.74}$$

In the particular case of cylindrical coordinates and spherical coordinates, respectively, this operator is given by:

$$\begin{aligned}
\Delta &= \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial z^2} \\
\Delta &= \frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2\frac{\partial}{\partial R}\right) + \frac{1}{R^2\sin\varphi}\frac{\partial}{\partial \varphi}\left(\sin\varphi\frac{\partial}{\partial \varphi}\right) + \frac{1}{R^2\sin^2\varphi}\frac{\partial^2}{\partial \theta^2}
\end{aligned} \tag{4.75, 76}$$

4.1.8 *Differentiation of Vectors in Curvilinear Coordinates*

4.1.8.1 Christoffel Symbols of the First and Second Kind

Relation (3.21) can be differentiated with respect to the variable q^k. The result is

$$\mathbf{e}_{i,k}\cdot\mathbf{e}_j + \mathbf{e}_i\cdot\mathbf{e}_{j,k} = g_{ij,k} \qquad (i,j,k = 1,2,3) \tag{4.77}$$

This expression can be rewritten by circular permutation:

$$\begin{aligned}
\mathbf{e}_{j,i}\cdot\mathbf{e}_k + \mathbf{e}_j\cdot\mathbf{e}_{k,i} &= g_{jk,i} \\
\mathbf{e}_{k,j}\cdot\mathbf{e}_i + \mathbf{e}_k\cdot\mathbf{e}_{i,j} &= g_{ki,j} \quad (i,j,k = 1,2,3)
\end{aligned} \tag{4.78}$$

By taking into account the relation (4.8) and the fact that the mixed second order derivatives of the position vector **r** do not depend on the differentiation order, it can be shown that:

$$\mathbf{e}_{i,j} = \mathbf{e}_{j,i} \tag{4.79}$$

The Christoffel symbols of first kind can be defined taking into account (4.79). First, the relations (4.78) are added. Next, the relationship (4.77) is subtracted. The result is:

$$[ij,k] \equiv \begin{bmatrix} ij \\ k \end{bmatrix} \equiv \Gamma_{kij} \equiv \mathbf{e}_{i,j} \cdot \mathbf{e}_k = \frac{1}{2}\left(-g_{ij,k} + g_{jk,i} + g_{ki,j}\right) \quad (i,j,k = 1,2,3) \tag{4.80}$$

Here three common notations have been used to designate the Christoffel symbols of first kind. The Christoffel symbols of the second kind are defined as follows:

$$\begin{Bmatrix} k \\ i\ j \end{Bmatrix} = \Gamma^k_{ij} = g^{kl}[ij,l] \quad (i,j,k = 1,2,3) \tag{4.81}$$

The index variables i and j in the Christoffel symbols are interchangeable, meaning that the next relations can be written:

$$[ij,k] = [ji,k], \quad \begin{Bmatrix} k \\ i\ j \end{Bmatrix} = \begin{Bmatrix} k \\ j\ i \end{Bmatrix} \quad (i,j,k = 1,2,3) \tag{4.82}$$

It follows that the Christoffel symbols of the first and the second kind, respectively, contain 18 distinct symbols each.

The Christoffel symbols are useful in calculating the derivatives of the elements of the fundamental forms. For example, the expression (4.81) is multiplied by g_{km} and the relationship (3.26) is used, resulting

$$[ij,k] = g_{kl}\begin{Bmatrix} l \\ i\ j \end{Bmatrix} \quad (i,j,k = 1,2,3) \tag{4.83}$$

Then, using Eq. (4.81) one finds that the Christoffel symbols are linked with the derivatives of the vectors of the reciprocal basis, i.e.

$$\begin{Bmatrix} k \\ i\ j \end{Bmatrix} = \mathbf{e}_{i,j} \cdot \mathbf{e}_k, \quad \mathbf{e}_{i,j} = \begin{Bmatrix} k \\ i\ j \end{Bmatrix}\mathbf{e}_k \quad (i,j,k = 1,2,3) \tag{4.84}$$

Further, by using relations (4.77) and (4.83), the following expression is obtained, in which derivatives of the components of the fundamental form occur:

$$g_{ij,k} = g_{jl}\left\{ {l \atop i\ k} \right\} + g_{il}\left\{ {l \atop j\ k} \right\} \quad (i,j,k = 1,2,3) \tag{4.85}$$

At the same time, the relationship (3.26) allows writing a useful expression:

$$g_{ik,l}g^{kj} + g_{ik}g^{kj}_{,i} = 0 \tag{4.86}$$

This expression leads to a relationship which is similar with Eq. (4.85), i.e.:

$$g^{ik}_{,k} = -g^{jk}\left\{ {i \atop l\ k} \right\} - g^{il}\left\{ {j \atop l\ k} \right\} \quad (i,j,k = 1,2,3) \tag{4.87}$$

Finally, one can show that

$$\frac{\partial}{\partial q^i}\ln(g)^{1/2} = \left\{ {j \atop i\ j} \right\} \quad (i = 1,2,3) \tag{4.88}$$

Note that this derivative with respect to the curvilinear coordinates is expressed in terms of the Christoffel symbols. To obtain the expression (4.88), the development of the determinant $g = \det\ |g_{ij}|$ with respect to the elements of a row (or of a column) has been used.

The transformation law of the Christoffel symbols of second kind is expressed as

$$\left\{ {k \atop i\ j} \right\}' = \left\{ {n \atop l\ m} \right\} \frac{\partial x^i}{\partial x'^i}\frac{\partial x^m}{\partial x'^j}\frac{\partial x'^k}{\partial x'^n} + \frac{\partial^2 x^p}{\partial x'^i \partial x'^j}\frac{\partial x'^k}{\partial x^p} \quad (i,j,k = 1,\ldots,n) \tag{4.89}$$

In the case that the Christoffel symbols of second kind are identically zero with respect to a coordinate system $\{x'^i\}$, the fundamental form g_{ij} has constant components h_{ij}. It can be shown that this is a necessary and sufficient condition.

4.1.8.2 Covariant Derivative

Derivative of a Scalar

Consider a scalar function U. The partial derivatives $U_{,i}$ constitute the covariant components of a vector. Indeed, if one takes into account the condition which must be met by a scalar function, the partial derivatives in a new reference frame, denoted $U'_{,j}$ are given by:

$$U'_{,j} = U_{,i}\frac{\partial q^i}{\partial q'^j} \quad (j = 1,2,3) \tag{4.90}$$

The covariant derivative of a scalar U of weight P may be written as

$$U_{;i} = U|_i = U_{,i} - P\left\{ \begin{matrix} j \\ i\ j \end{matrix} \right\} U \quad (i = 1, 2, 3) \tag{4.91}$$

Derivative of a Vector

Considered a vector $\mathbf{V}$. Taking into consideration the representation (3.9), the partial derivative $\mathbf{V}_{,j}$ of this vector can be expressed as a function of the contravariant components of the vector and of its derivatives, in the following way:

$$\mathbf{V}_{,j} = V^i_{,j}\mathbf{e}_i + V^i\mathbf{e}_{i,j} \tag{4.92}$$

The covariant derivative of the contravariant components of a vector are defined as follows

$$V^i_{;j} \equiv V^i|_j = V^i_{,j} + \left\{ \begin{matrix} i \\ j\ k \end{matrix} \right\} V^k \quad (i,j = 1, 2, 3) \tag{4.93}$$

Using this definition, the previous relation can be written in a more compact way:

$$\mathbf{V}_{,j} = V^i_{;j}\mathbf{e}_i \quad (j = 1, 2, 3) \tag{4.94}$$

Similarly, the covariant derivative of the covariant components of the vector can be defined by the relationship:

$$V_{i;j} \equiv V_i|_j = V_{i,j} - \left\{ \begin{matrix} k \\ i\ j \end{matrix} \right\} V_k \quad (i,j = 1, 2, 3) \tag{4.95}$$

Thus, the partial derivative $\mathbf{V}_{,j}$ of the vector is given by:

$$V_{,j} = V_{i;j}\mathbf{e}^i \quad (j = 1, 2, 3) \tag{4.96}$$

The two covariant derivatives defined above allow the differentiation of the components of a vector. Conversely, if one takes into account the relationships (4.94) and (4.96), it is obvious that both the basis $\{\mathbf{e}_i\}$ and the reciprocal basis $\{\mathbf{e}^j\}$ remain apparently constant.

Differential of a Vector

Two differentials of a vector are defined: the differential of the contravariant components and the differential of the covariant components, respectively:

$$DV^i \equiv dV^i + \left\{ \begin{matrix} i \\ j\ k \end{matrix} \right\} V^k dq^j, \quad DV_i \equiv dV_i - \left\{ \begin{matrix} k \\ i\ j \end{matrix} \right\} V_k dq \quad (i = 1, 2, 3) \qquad (4.97, 98)$$

Thus, the differential of the vector can be expressed in two ways, as follows:

$$d\mathbf{V} = (DV^i)\mathbf{e}_i = (DV_j)\mathbf{e}^j \qquad (4.99)$$

Particular Cases

If all components g_{ij} of the fundamental form are constant, then all the Christoffel symbols are zero. Therefore, the covariant differentiation is reduced to ordinary differentiation. Note that, in general, the covariant derivative of a vector with constant components is not null.

The case of orthogonal Cartesian coordinate systems is often encountered in practice. It is characterized by $g_{ij} = \delta_{ij}$ and, according to the above theory, in this coordinate system, the covariant derivative reduces to the usual derivative. The formalism of modern physics requires that the vector relationships are invariant to the change of the coordinate system. This means that the equalities valid in orthogonal Cartesian coordinate systems must be valid into an arbitrary coordinate system. Consequently, to the operator of ordinary differentiation in a Cartesian coordinate system must correspond the operator of covariant differentiation in the other coordinate system. This is achieved by a transcription which moves in a convenient way the index variables.

For example, consider the divergence of a vector **V** and the Laplace operator. If one takes into account the general relationship (4.88), the expression of the divergence can be written as:

$$\operatorname{div} \mathbf{V} = \frac{1}{g^{1/2}} \left(g^{1/2} V^i \right)_{,i} \qquad (4.100)$$

If the vector is conservative, i.e. $\mathbf{V} = grad\ U$, it follows that $V^i = g^{ij} U_{,j}$. The Laplace operator applied to the scalar U can be written as:

$$\Delta U = \operatorname{div} \operatorname{grad} U = \frac{1}{g^{1/2}} \left(g^{ij} g^{1/2} U_{,j} \right)_{,i} \qquad (4.101)$$

In an orthogonal Cartesian coordinate system, the divergence has the form shown in the relationship (3.68). By using the covariant differentiation, one obtains a formally simpler expression:

$$\operatorname{div} \mathbf{V} = V^i_{;i} \qquad (4.102)$$

4.1.9 Intrinsic Derivative

Consider a domain $D \subset R^3$ and a one-dimensional manifold (a curve) C in this domain. The curve is defined parametrically, by using the bounded parameter $t \in [t_1, t_2]$ and the differentiable relations $C : x^i = x^i(t) \quad (i = 1, 2, 3)$. Consider a differentiable vector field $\mathbf{V} = \mathbf{V}(x^1, x^2, x^3)$, defined in the domain D. The vectors $\mathbf{V}$ which are defined on the one-dimensional manifold C depend of parameter t. Then, the following relationship occurs

$$\frac{d\mathbf{V}}{dt} = \mathbf{V}_{,j}\dot{x}^j = V^i_{;j}\dot{x}^j\mathbf{e}_i \tag{4.103}$$

If one takes into account the expression (4.93) of the covariant derivative, the relationship (4.103) can be rewritten as follows

$$\frac{d\mathbf{V}}{dt} = \mathbf{V}_{,j}\dot{x}^j = \frac{\partial V^i}{\partial t} + \left\{ \begin{matrix} i \\ j \quad k \end{matrix} \right\} V^j x^k \tag{4.104}$$

where the following notation has been used:

$$\frac{\partial V^i}{\partial t} \equiv V^i_{,j}\dot{x}^j \quad (i = 1, 2, 3) \tag{4.105}$$

Relation (4.104) is the *intrinsic derivative* (or the *absolute derivative*) of the contravariant components V^i of the vector, with respect to the parameter t. The intrinsic differentiation keeps the usual differentiation rules.

4.1.10 Parallel Transport

Consider a one-dimensional manifold (a curve) C of parameter t and a point P on that curve. In this point, consider a vector $\mathbf{V}$ of arbitrary orientation. The point P is moving on the curve, yielding a field of parallel vectors along the curve, and this field does not depend of the parameter t. For this reason, on the whole curve C the following simple relationship is fulfilled:

$$\frac{d\mathbf{V}}{dt} = 0 \tag{4.106}$$

Therefore, the contravariant components of the vector $\mathbf{V}$ satisfy the differential equations system

$$\frac{\partial V^i}{\partial t} + \left\{ \begin{matrix} i \\ j \;\; k \end{matrix} \right\} V^j \dot{x}^k = 0 \quad (i = 1, 2, 3) \tag{4.107}$$

Consider a three-dimensional Euclidean manifold, for which the concept of field of parallel vectors makes sense. Equation (4.107) can be rewritten as

$$\left(V^i_{,k} + \left\{ \begin{matrix} i \\ j \;\; k \end{matrix} \right\} V^j \right) \dot{x}^k = 0 \quad (i = 1, 2, 3) \tag{4.108}$$

Note that, for a given point, this relationship is checked for any curve C passing through that point. The consequence is that any field of parallel vectors in the Euclidean space E_3 fulfill the system of equations

$$V^i_{;k} = V^i_{,k} + \left\{ \begin{matrix} i \\ j \;\; k \end{matrix} \right\} V^j = 0 \quad (i, k = 1, 2, 3) \tag{4.109}$$

Relationships (4.109) are the so-called parallel transport conditions of the contravariant components of the vector **V**. In the case of the covariant components of the vector **V**, the conditions for parallel transport are

$$V_{i;k} = V_{i,k} + \left\{ \begin{matrix} j \\ i \, k \end{matrix} \right\} V_j = 0 \quad (i, k = 1, 2, 3) \tag{4.110}$$

Consider a curve C depending on the bounded curvilinear coordinate $s \in [s_1, s_2]$. The curve C is defined by the equations $x^i = x^i(s) \;\; (i = 1, 2, 3)$. Assume that a field of parallel vectors **V** exists on the curve C. Then, the following relationships should be fulfilled:

$$\frac{\partial V^i}{\partial s} + \left\{ \begin{matrix} i \\ j \, k \end{matrix} \right\} V^j \frac{dx^k}{ds} = 0 \quad (i = 1, 2, 3) \tag{4.111}$$

Previously, parallel vectors of arbitrary direction with respect to the curve C have been considered. Now, consider a field of vectors which are tangent to that curve. The field is described by the quantities $dx^i/ds \;\; (i = 1, 2, 3)$. Using the relation (4.111), which represents the condition for those vectors to be parallel each other, one obtains

$$\frac{d^2 x^i}{ds^2} + \left\{ \begin{matrix} i \\ j \, k \end{matrix} \right\} \frac{dx^j}{ds} \frac{dx^k}{ds} = 0 \quad (i = 1, 2, 3) \tag{4.112}$$

These relationships constitute the equations of a straight line in a curvilinear coordinate system. Consider now the particular case of an orthogonal Cartesian coordinate system. In this situation, the Christoffel symbols are canceled and the relationships (4.112) reduce to $d^2 x^i / ds^2 = 0 (i = 1, 2, 3)$.

The notions can be extended to the general case of a Riemannian manifold V_n of metric (4.1). In this case, the length of the arc of curve is given by

$$s = \int_{t_1}^{t_2} (g_{ij}x^i x^j)^{1/2} dt \tag{4.113}$$

The geodesics of the manifold V_n are given by the extremal curves of the functional (4.113). To find these extremal curves, one may apply the known results of variational calculation (Badescu 2003). Thus, starting from the corresponding Euler–Lagrange equations, one obtains, after some processing, the following differential equations:

$$\frac{d^2x^i}{ds^2} + \left\{ \begin{matrix} i \\ j\ k \end{matrix} \right\} \frac{dx^j}{ds}\frac{dx^k}{ds} = 0 \quad (i = 1, \ldots, n) \tag{4.114}$$

These are the equations of the extremal curves of the functional (4.113) and, at the same time, the equations of the geodesic lines in the manifold V_n. Note that Eq. (4.114) have the same form as the Eq. (4.112) of the straight line in the Euclidean space E_3. Some examples: on a sphere, the geodesics are great circles and on a cylinder or on a cone, the geodesics are the generators of the cylinder or the cone, respectively.

Consider the particular case of a Euclidean manifold V_n. There is a coordinates system where the Christoffel symbols are zero. In that system, the relations (4.114) reduce to the following relationships, showing that geodesics in the Euclidean space are straight lines:

$$\frac{d^2x^i}{ds^2} = 0 \quad (i = 1, \ldots, n) \tag{4.115}$$

Finally, consider the case of a curve on the manifold V_n. The previous theory can be extended to build parallel vectors along that curve. It can be shown that the vectors obtained by the parallel transport of a vector which is tangent to a geodesic, are permanently tangent to that geodesic. It is said that the geodesics are *self-parallel curves*.

4.2 Recent Formalizations of Riemann Geometry

In this section, concepts and notations used recently in the presentation of Riemann geometry are briefly present. The exposure of Casetti et al. (2000) will be closely followed. The main purpose of the section is to create an intuitive image about the main concepts, since the approach has not the purpose to constitute a rigorous treatment of the subject. For a more elaborated introduction the reader is sent to the general relativity book of Landau and Lifshitz (2000) or to more complete works, but still easily accessible, such as Wald (1984) and Ionescu-Pallas (1980). A more rigorous mathematical treatment, but accessible to physicists and engineers is given by Carmo (1993). Finally, an exhaustive and rigorous presentation, but far exceeding the level of this book is the work of Kobayashi and Nomizu (1991).

In this section, the Einstein summation convention for the dummy indices is systematically used. The dependence on time of the vector and tensor components is not explicitly mentioned. This will be done only when necessary.

4.2.1 Manifolds

In a space of dimension m, one can define a hypersurface of dimension m (or, in other words, a m-hypersurface) by specifying an equation of the form:

$$g(x_1, \ldots, x_m) = 0 \tag{4.116}$$

It is said that the hypersurface is *smooth* if the function g is differentiable and the vector of the normal at the surface (4.116), given by the following expression, where the gradient operator appears:

$$\text{grad } g \equiv \left\{ \frac{\partial g}{\partial x_1}, \ldots, \frac{\partial g}{\partial x_m} \right\} \tag{4.117}$$

does not cancel. Consider r hypersurfaces of dimension m, given by the system of equations:

$$\begin{aligned} g_1(x_1, \ldots, x_m) &= 0 \\ \cdots & \\ g_r(x_1, \ldots, x_m) &= 0 \end{aligned} \tag{4.118}$$

The intersection M of all these hypersurfaces, i.e. the set of all points that satisfy the Eq. (4.118), is called manifold of dimension m–r, provided that the vectors of the normals to the r hypersurfaces, grad $g_1, \ldots,$ grad g_r, are linearly independent at any point.

Thus, a manifold z-dimensional is given by m–z equations. Also, a manifold of dimension $m - 1$ is conceived as a hypersurface.

4.2.1.1 Riemannian Manifolds

A set of coordinates, i.e. a set of real numbers $(x^1, \ldots, x^n)$ representing the "position" of a point on a manifold, is called *chart*. A collection of charts on a manifold is called *atlas* on that manifold. A set M is called *differentiable manifold* if it can be covered with a collection, finite or numerable, of charts, so that each point of M is represented in at least one chart, and different charts are differentially connected to each other. The number n of coordinates of the chart is the same, for each connected part the manifold (and for the whole manifold, if it is connected, i.e. if it cannot be separated into two disjoint parts which still preserves the property that they are manifolds). Such a number is called the *dimension of the manifold*.

Vectors and Tensors

A vector (more precisely, a *tangent vector*) can be defined by using curves on the manifold M. Given a curve γ on M, represented in local coordinates by the parametric equations $x = \varphi(t)$, a tangent vector is defined in the point $P \in M$ as the velocity vector of the curve in the point P, i.e.:

$$v = \dot{\gamma} = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t}, \quad \varphi(0) = P \tag{4.119}$$

so that the n components of the tangent vector v are given by:

$$v^i = \frac{d\varphi^i}{dt} \tag{4.120}$$

The set of all tangent vectors of M in P forms a linear space called *tangent space* of M in P, which is denoted $T_P M$. Each tangent space is isomorphic with the n-dimensional Euclidean space. Given a chart $(x^1, \ldots, x^n)$ in the neighborhood of P, a basis $(X_1, \ldots, X_n)$ of $T_P M$ can be defined, so that an arbitrary vector v is expressed as the sum of the vectors X_i of the basis, weighted by its own components:

$$v = v^i X_i \tag{4.121}$$

The basis $\{X_i\}$ is called basis of coordinates of $T_P M$ and its components X_i are usually denoted by $\partial / \partial x^i$. The origin of this notation comes from the fact that the tangent vectors can be defined as directional derivatives on M (see, for example, Wald 1984). The basis depends on the chart. By choosing a different chart,$(x'^1, \ldots, x'^n)$, one obtains another basis, $\{X'_i\}$. The components of v in the two bases are related through the rule:

$$v'^i = v^j \frac{\partial x'^i}{\partial x^j} \tag{4.122}$$

which is called the *vectors transformation rule*. Relationship (4.122) may serve as a definition of the tangent vector, which may be seen as a set of numbers whose components transform themselves according to that relation. The set of all tangent spaces on the manifold M, denoted

$$TM = \bigcup_{P \in M} T_P M \tag{4.123}$$

is a $2n$-dimensional manifold, called the *tangent bundle* of M.

A vector field on the manifold M is obtained by attaching a vector v_P to each point $P \in M$. If f is a smooth function, then, for any $P \in M$, the result

$$V(f)|_P = v_P(f) \tag{4.124}$$

is a real number; in other words $v(f)$ is a function defined on M. If $v(f)$ is a smooth function, one says that V is a *smooth vector field* on M. The curves $\varphi(t)$ which satisfy the differential equations

$$\dot{\varphi} = V(\varphi(t)) \tag{4.125}$$

are called *trajectories* of the field V, and the map $\varphi_t : M \to M$ defined for any point P of M along the trajectory of V passing through the P is called *flow* of V. Given two vector fields V and W, their *commutator* is defined as a vector field $[V, W]$ so that

$$[V, W](f) \equiv V(W(f)) - W(V(f)) \tag{4.126}$$

or, using local components,

$$[V, W]^j = V^i \frac{\partial W^j}{\partial x^i} - W^i \frac{\partial V^j}{\partial x^i} \tag{4.127}$$

Note that if $\{X_i\}$ is a coordinate basis, then the following relation can be written

$$\left[X_i, X_j\right] = 0 \qquad \forall i, j \tag{4.128}$$

and that, conversely, giving n vector fields that do not cancel, do commute and are linearly independent, then there is always a chart for which these vectors constitute a coordinate basis.

The tangent vectors are not the only vector-like quantities that can be defined on the manifold M. There are also *cotangent vectors* that can be defined as follows. First, remind that the *dual space* V^* of a vector space V is the space of linear maps from V in the set of real numbers. Given a basis of V, denoted $\{u_i\}$, then a basis of V^*, noted $\{u*^i\}$, called *dual basis*, is defined by

$$u *^i \left(u_j\right) = \delta^i_j \tag{4.129}$$

The dual space of *TM*, denoted *T*M*, is called *cotangent bundle* of *M*. Its elements are called cotangent vectors, or *covariant vectors* (while the tangent vectors are called *contravariant vectors*). The elements of the dual basis are usually denoted $dx^1, \ldots, dx^n$ and dx^i has the property that $dx^i(\partial/\partial x^j) = \delta^i_j$. The components ω_i of the cotangent vectors transform themselves by the rule

$$\omega'_i = \omega_j \frac{\partial x^j}{\partial x'^i} \tag{4.130}$$

This relationship can be compared with Eq. (4.122). The usual rule is to use subscripts to designate the components of the dual vectors and superscripts to designate the components of vector.

A (*k*, *l*)-*tensor* over the vector space *V* is an multi-linear map

$$T : \left(\underbrace{V* \times \cdots \times V*}_{k \text{ times}}\right) \times \left(\underbrace{V \times \cdots \times V}_{l \text{ times}}\right) \to R \tag{4.131}$$

i.e. acting on *k* dual vectors and *l* vectors, *T* produces a real number, and does this in a way that, if one fixes all the vectors or dual vectors, except one, it is a linear map in the remaining variable. With this notation, a (0, 0) tensor is a scalar, a (0, 1) tensor is a vector and a (1, 0) tensor is a dual vector. The space $\tau(k, l)$ of tensors of type (*k*, *l*) is a linear space. A (*k, l*)-tensor is defined once its action on the *k* vectors of the dual basis and the *l* vectors of the base is known. Since there are $n^k n^l$ independent ways to choose these basis vectors of the two bases, $\tau(k, l)$ is a n^{k+l}-dimensional linear space.

On the set of tensors one can define two operations. The first one is called *contraction* with respect to the arguments of the *i*th dual vector and *j*th vector, respectively, being a map of the type

$$C : T \in \tau(k, l) \to CT \in \tau(k-1, l-1) \tag{4.132}$$

defined by

$$CT = \sum_{\sigma=1}^{n} T\left(\ldots, \underbrace{v*^{\sigma}}_{i}, \ldots; \ldots, \underbrace{v^{\sigma}}_{j}, \ldots\right) \tag{4.133}$$

The contracted tensor CT is independent of the choice of the basis, so that the contraction is a well defined, invariant, operation. The second operation is the *tensor product*, which maps an element $\tau(k, l) \times \tau(k', l')$ in the element $\tau(k + k', l + l')$, i.e., it maps two tensors *T* and *T'* in a new tensor, denoted by $T \otimes T'$, defined as follows: being given *k* + *k*' dual vectors $v*^1, \ldots, v*^{k+k'}$ and *l* + *l*' vectors $w_1, \ldots, w_{l+l'}$, then

$$\begin{aligned} & T \otimes T'\left(v*^1, \ldots, v*^{k+k'}; w_1, \ldots, w_{l+l'}\right) \\ & \quad = T\left(v*^1, \ldots, v*^k; w_1, \ldots, w_l\right) T\left(v*^{k+1}, \ldots, v*^{k+k'}; w_{l+1}, \ldots, w_{l+l'}\right) \end{aligned} \tag{4.134}$$

The tensor product enables to construct a basis for $\tau(k,l)$ starting from a basis $\{v_\mu\}$ of V and its dual basis $\{v*^\nu\}$; such a base is given by the n^{k+l} tensors $\{v_{\mu_1} \otimes \cdots \otimes v_{\mu_k} \otimes v*^{\nu_1} \otimes \cdots \otimes v*^{\nu_l}\}$. Thus, every tensor $T \in \tau(k,l)$ allows the decomposition

$$T = \sum_{\mu_1,\ldots,\nu_l=1}^{n} T^{\mu_1 \cdots \mu}_{\nu_1 \ldots \nu_l} v_{\mu_1} \otimes \cdots \otimes v*^{\nu_l} \tag{4.135}$$

where the numbers $T^{\mu_1 \cdots \mu}_{\nu_1 \ldots \nu_l}$ are called the components of T in basis $\{v_\mu\}$. The components of the contracted tensor CT are

$$(CT)^{\mu_1 \cdots \mu_{k-1}}_{\nu_1 \ldots \nu_{l-1}} = T^{\mu_1 \ldots \sigma \ldots \mu_k}_{\nu_1 \ldots \sigma \ldots \nu_l} \tag{4.136}$$

(one reminds that the Einstein summation convention is used), and the components of the tensor product $T \otimes T'$ are

$$(T \otimes T')^{\mu_1 \cdots \mu_{k+k'}}_{\nu_1 \ldots \nu_{l+l'}} = T^{\mu_1 \cdots \mu_k}_{\nu_1 \ldots \nu_l} T'^{\mu_{k+1} \cdots \mu_{k+k'}}_{\nu_{l+1} \cdots \nu_{l+l'}} \tag{4.137}$$

All these results are valid for an arbitrary vector space. Therefore, they hold for the particular cases of the vector spaces of the tangent bundle *TM* of *M*. Over these spaces, tensors and tensor fields can be defined (analogously to the vector fields).

Riemann Metrics

The square of the infinitesimal distance on the manifold *M*, i.e. the length of the element ds^2 (which is often called *the metric*) can be defined at any point $P \in M$ by using a (0,2)—tensor g, provided that it is symmetric, meaning $g(v,w) = g(w,v)$ and non-degenerate, i.e. $g(v,w) = 0$, $\forall v \in T_P M$, if and only if $w = 0$. Considering that g is a symmetric tensor, the last condition may be rephrased in the following way: $g(v,w) = 0$, $\forall v \in T_P M$ if and only if $v = 0$. In fact, a tensor g with these properties induce on the tangent bundle *TM* a non-degenerate quadratic form (called *scalar product*), which is denoted as follows:

$$g(v,w) \equiv \; <v,w> \; : TM \times TM \to R \tag{4.138}$$

The introduction of this metric makes it possible to measure the lengths on that manifold. A manifold M, equipped with a scalar product is called *Riemannian manifold* and the scalar product thus defined is considered a *Riemannian structure* on the manifold M. If the quadratic form (4.138) is positive definite, it is said that it is *Riemann metric*. In this case, the square of the length element is always positive. The length of a curve can be defined as follows:

$$l(\gamma) = \int_{\gamma} (<\dot{\gamma}, \dot{\gamma}>)^{1/2} \tag{4.139}$$

The curves (denoted γ) that are extremals of the length functional are called *geodesics* of M.

In a coordinate basis, the metric g can be expanded as follows:

$$g = g_{ij} dx^i \otimes dx^j \tag{4.140}$$

Therefore, in local coordinates one can define (the square of) the length element, invariant on the manifold, in the following way:

$$ds^2 = g_{ij} dx^i dx^j \tag{4.141}$$

The scalar product of two vectors v and w is given, in terms of g, by

$$<v, w> \ = g_{ij} v^i w^j = v_j w^j = v^i w_i \tag{4.142}$$

In the above relationship, it has been made use of the fact that g makes a bi-univocal correspondence between the vectors and the dual vectors, which, in components, looks like this:

$$g_{ij} v^j = v_i \tag{4.143}$$

For this reason, the components of the inverse metric g^{-1} are simply denoted g^{ij} instead of $(g^{-1})^{ij}$. The inverse metric allows the passage from the (covariant) components of the dual vector to the (contravariant) components of the vector:

$$g^{ij} v_j = v^i \tag{4.144}$$

The operation of raising and lowering the indices can be applied not only to the vector components, but also to the components of tensors. This allows the passage from the components of a tensor (k, l) to the corresponding components of the tensor $(k + 1, l - 1)$, and vice versa. Doing this, does not change the sum $k + l$, which is called the *rank* (or the *order*) of the tensor.

4.2.2 Covariant Differentiation

As seen, the introduction of a differential calculus on a manifold which is not Euclidean is complicated by the fact that ordinary derivatives do not map common vectors into vectors. This means that, for example, the ordinary derivatives of the components of a vector w, denoted dw^i/dt, calculated at a point P on the given curve $\gamma(t)$, do not represent the components of a vector in T_PM, because they do not transform themselves according to the rule (4.122). The geometric explanation of this fact is that the parallel transport of a vector from the point P to a point Q in a non-Euclidean manifold depends on the chosen path to go from P in Q. Indeed, to define the derivative of a vector in P, first, this vector must be moved from P in a neighboring point Q on the curve. To evaluate the difference between the vector in P and the vector in Q, the vector must be parallel transported back to the point P. It follows that the rigorous definition of the differentiation operation requires a clear definition of what the parallel transport is. Conversely, if it has been already defined a consistent differentiation operation, i.e. a differentiation operation which maps vectors into vectors, the parallel transport can be defined by the statement that a vector is parallel transported along a curve if its derivative along that curve is null. The two procedures are equivalent. Next, the first procedure will be followed, by introducing the notion of *connection*, and using it to define the differentiation operation. Such a derivative is called covariant derivative.

Connections can be defined by using tensor fields but the definition using vectors is simpler. A (linear) connection on a manifold M is a map that, from two vector fields A and B, yields a third vector field, labeled $\nabla_A B$, with the following properties:

1. $\nabla_A B$ is bilinear in A and B, meaning that $\nabla_A(\alpha B + \beta C) = \alpha \nabla_A B + \beta \nabla_A C$ and $\nabla_{\alpha A + \beta B} C = \alpha \nabla_A C + \beta \nabla_B C$;
2. $\nabla_{f(A)} B = f(\nabla_A B)$;
3. $\nabla_A f(B) = (\partial_A f)B + f(\nabla_A B)$, where ∂_A is the ordinary directional derivative in the direction of A. This last property is called *Leibnitz rule*.

One defines the *parallel transport* of a vector V along a curve γ, whose tangent vector field is denoted $\dot{\gamma}$, as the (unique) vector field $W(t) = W(\gamma(t))$ along $\gamma(t)$ having the following properties

1. $W(0) = V$;
2. $\nabla_{\dot{\gamma}} W = 0$ along γ.

The notion of covariant derivative immediately follows: the *covariant derivative* DV/dt of V along γ is given by the vector field

$$\frac{DV}{dt} = \nabla_{\dot{\gamma}} V \tag{4.145}$$

Based on the Eq. (4.145) (and forcing the language) it is often said about $\nabla_X Y$ that it is the covariant derivative of Y along X, where X and Y are two arbitrary vector fields. For a given metric g, among all possible linear connections, there is one and only one, with the following properties:

(i) it is symmetric, i.e.

$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{4.146}$$

(ii) it preserves the scalar product, i.e. the scalar product of two parallel vector fields, P and P', is constant along γ, according with the equation

$$\frac{d}{dt} < P, P' > \; = 0 \tag{4.147}$$

Such a linear connection is obviously the natural one on a Riemannian manifold, being called *Levi-Civita connection* (or *Riemannian connection*). Further usage of the covariant derivative without prior specifications implicitly assumes that it is the covariant derivative induced by the Riemannian connection.

The components of the Riemannian connection ∇ with respect to a basis of coordinates $\{X_i\}$ are called *Christoffel symbols of the second kind*, given by

$$\Gamma^i_{jk} = \; < dq^i, \nabla_{X_j} X_k > \tag{4.148}$$

These symbols can be expressed in terms of the derivatives of the metric components, following the relationship:

$$\Gamma^i_{jk} = \frac{1}{2} g^{im} \left(\partial_j g_{km} + \partial_k g_{mj} - \partial_m g_{jk} \right) \tag{4.149}$$

where $\partial_i = \partial / \partial x^i$. In local coordinates, the expression (4.145) of the covariant derivative of a vector field V is

$$\frac{DV^i}{dt} = \frac{dV^i}{dt} + \Gamma^i_{jk} \frac{dx^j}{dt} V^k \tag{4.150}$$

4.2.2.1 Geodesics

Geodesic curves have been defined as curves of extremal length on a manifold. They can also be defined as *self-parallel curves*, i.e. curves having the property that the tangent vector is always parallel transported. Consequently, the geodesics are those curves $\gamma(t)$ that satisfy the following equation (called *geodesic equation*):

$$\frac{D\dot{\gamma}}{dt} = 0 \tag{4.151}$$

In local coordinates the expression of this equation is obtained from Eq. (4.150):

$$\frac{d^2x^i}{dt^2} + \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0 \tag{4.152}$$

Since the norm of the tangent vector $\dot{\gamma}$ of a geodesic is constant, $|d\gamma/dt| = c$, the arc length on a geodesic is proportional to the parameter

$$s(t) = \int_{t_2}^{t_2} |\frac{d\gamma}{dt}|dt = c(t_2 - t_1) \tag{4.153}$$

If the parameter is actually the arc length, i.e. $c = 1$, it is said that the geodesic is normalized. In the following, when a geodesic is considered, it is implicitly assumed that it is normalized. This means that Eq. (4.152) are nothing but *the Euler–Lagrange equations* for the length functional on the curve $\gamma(s)$ parameterized by the arc length

$$l(\gamma) = \int_{\gamma} ds \tag{4.154}$$

Given a geodesic $\gamma(s)$ on M, there is only one vector field G on TM so that its trajectories are $(\gamma(s), \dot{\gamma}(s))$. Such a vector field is called *geodesic field* and its flow is called *geodesic flow* on M.

4.2.3 Curvature

Intuitively, the curvature of a Riemannian manifold (M, g) is a way to measure how much this manifold differs from an Euclidean manifold. The *curvature tensor*, which is also called *Riemann-Christoffel tensor*, is a tensor of order four defined as follows:

$$R(X, Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]} \tag{4.155}$$

where ∇ is the Levi-Civita connection on M. One sees that, if $M = R^N$, then $R(X, Y) = 0$ for all pairs of tangent vectors X, Y, due to the commutativity of the ordinary derivatives. In addition, R represents a measure of the non-commutativity of the covariant derivative. In fact, once the coordinate system $\{x_1, \ldots, x_n\}$ is chosen, one finds (since $\left[\partial/\partial x_i, \partial/\partial x_j\right] = 0$):

$$R\left(\frac{\partial}{\partial x_i},\frac{\partial}{\partial x_j}\right) = \nabla_{\partial/\partial x_i}\nabla_{\partial/\partial x_j} - \nabla_{\partial/\partial x_j}\nabla_{\partial/\partial x_i} \tag{4.156}$$

In local coordinates, the components of the Riemann curvature tensor (which is a (1, 3)-tensor) are given by:

$$R^i_{jkl} = \frac{\partial\Gamma^i_{jl}}{\partial x^k} - \frac{\partial\Gamma^i_{kl}}{\partial x^j} + \Gamma^r_{jl}\Gamma^i_{kr} - \Gamma^r_{kl}\Gamma^i_{jr} \tag{4.157}$$

For a given metric g, the curvature R is uniquely defined. A manifold (M, g) is said to be flat when the curvature tensor cancels.

Given a positive function f^2, the transformation specified by the following equation

$$(M,g) \to (M,\bar{g}); \quad \bar{g} \equiv f^2 g \tag{4.158}$$

is called *conformal transformation*. Two Riemann manifolds are said to be *conformally related* if they are linked by a conformal transformation. In particular, a manifold (M, g) is *conformally flat* if it is possible to find a conformal transformation that transform g into a flat metric. The conformally flat manifolds are characterized by significant simplification of the calculation of the curvature tensor components.

Close to the notion of curvature tensor is the sectional curvature (or Riemann curvature), which is defined next. Consider two vectors $u, v \in T_P M$, and defines

$$|u \wedge v| \equiv \left(|u|^2 \cdot |v|^2 - <u,v> \right)^{1/2} \tag{4.159}$$

which is the area of a two-dimensional parallelogram determined by u and v. If $|u \wedge v| \neq 0$, then the vectors u, v determine a bi-dimensional subspace $\pi \subset T_P M$. The *sectional curvature* in point P relative to π is defined as follows:

$$K(\mathrm{P}; u,v) \equiv K(\mathrm{P},\pi) = \frac{<R(v,u)u,v>}{|u \wedge v|^2} \tag{4.160}$$

It can shown that the sectional curvature is independent of the choice of the two vectors $u, v \in \pi$. In local coordinates, Eq. (4.160) becomes

$$K(P; u,v) = R_{ijkl}\frac{u^i v^j u^k v^l}{|u \wedge v|^2} \tag{4.161}$$

Knowledge of K for the $N(N-1)$ planes generated by a maximal set of linearly independent vectors completely determines the curvature tensor R in point P. If $\dim(M) = 2$, then K will coincide with the *Gaussian curvature* of the surface, i.e. with the product of the reciprocals of the two curvature radii.

A manifold is said to be isotropic if $K(\mathrm{P},\pi)$ does not depend on the choice of the plane π. A remarkable result (*Schur's theorem*) states that in this case the curvature K is constant, i.e. it is no longer dependent on the point P.

Very important are certain ways of averaging the sectional curvature. A first example is the *Ricci curvature* K_R in P in the direction v, which is defined as the sum of the sectional curvatures in P relative to the planes determined by v and the $N-1$ directions orthogonal to v. More specifically, if $\{e_1,\ldots,e_{N-1},v=e_N\}$ is an orthogonal basis on T_PM and π_i is the plane generated by v and e_i, then

$$K_R(\mathrm{P},v)=\sum_{i=1}^{N-1}K(\mathrm{P},\pi_i) \tag{4.162}$$

A second example is the *scalar curvature* $\Re$ in P, given by the sum of the N Ricci curvatures in P,

$$\Re(\mathrm{P})=\sum_{i=1}^{N}K_R(\mathrm{P},e_i) \tag{4.163}$$

Such curvatures can be defined in terms of the components of the curvature tensor, as follows (in formulas, the dependence on the point P is implicit, since these components are local quantities). First, the *Ricci tensor* is defined as being a second order tensor whose components, R_{ij}, are obtained by contracting the first and the third indices in the Riemann tensor

$$R_{ij}=R^k_{ikj} \tag{4.164}$$

So

$$K_R(v)=R_{ij}v^iv^j \tag{4.165}$$

The right-hand side of the Eq. (4.165) is called "*the saturation*" of R_{ij} with respect to v. The scalar curvature can be obtained as the trace of the Ricci tensor, according to the relationship

$$\Re=R^i_i \tag{4.166}$$

In the case of a manifold of *constant curvature*, the components of the Riemann curvature tensor have the remarkably simple form

$$R_{ijkl}=K\left(g_{ik}g_{kl}-g_{il}g_{jk}\right) \tag{4.167}$$

where K is the constant sectional curvature, so that the Ricci tensor components are

$$R_{ij} = Kg_{ij} \tag{4.168}$$

and all the curvatures defined above are constant, being related through the following relationships, which are particularly useful in applications because they depend only on the parameter N:

$$K = \frac{1}{N-1} K_R = \frac{1}{N(N-1)} \Re \tag{4.169}$$

4.2.4 Jacobi Equation

In this subsection a way of deducing of the *Jacobi equation* is presented. The procedure is as follows. First, the geodesic separation vector field J is defined. Then, it is shown that J is actually a *Jacobi field*, i.e. it obeys the Jacobi equation.

A *geodesic congruence* is defined as a family of geodesics $\{\gamma_\tau(s) = \gamma(s, \tau) | \tau \in R\}$ passing through the neighboring V of a point of a manifold, smoothly depending on the parameter τ. A reference geodesic $\gamma(s, \tau_0)$ is chosen. Denote by $\dot{\gamma}(s)$ the tangent vector field to $\gamma(s, \tau_0)$ in s, i.e. the velocity vector field, whose components are

$$\dot{\gamma}^i = \frac{dx^i}{ds} \tag{4.170}$$

and denote by $J(s)$ the tangent vector field in τ_0 to the curves $\gamma_s(\tau)$ for a fixed s, i.e. a vector field of components

$$J^i = \frac{dx^i}{d\tau} \tag{4.171}$$

The field J is called *geodesic separation field*; it is a measure of the distance between nearby geodesics.

The demonstration that J is a Jacobi field follows. First, notice that the field J commutes with $\dot{\gamma}$, i.e. $[\dot{\gamma}, J] = 0$. In fact, from the definition of the commutator (Eq. (4.127)), the definition of J (Eq. (4.171)) and the definition of $\dot{\gamma}$ (Eq. (4.170)), one finds

$$[\dot{\gamma}, J]^i = \dot{\gamma}^j \frac{\partial J^i}{\partial x^j} - J^j \frac{\partial \dot{\gamma}^i}{\partial x^j} = \frac{\partial x^j}{\partial s} \frac{\partial J^i}{\partial x^j} - \frac{\partial x^j}{\partial \tau} \frac{\partial \dot{\gamma}^i}{\partial x^j} = \frac{\partial J^i}{\partial s} - \frac{\partial \dot{\gamma}^i}{\partial \tau} \tag{4.172}$$

Using again the Eqs. (4.171) and (4.170), one finds that

$$\frac{\partial J^i}{\partial s} = \frac{\partial}{\partial s}\frac{\partial x^i}{\partial \tau} = \frac{\partial}{\partial \tau}\frac{\partial x^i}{\partial s} = \frac{\partial \dot{\gamma}}{d\tau} \tag{4.173}$$

so $[\dot{\gamma}, J] = 0$. Now, the covariant derivative $\nabla_{\dot{\gamma}}^2 J$ of the field J is computed. Note that the covariant derivative comes from a Levi-Civita connection, which is symmetric (see Eq. (4.146)), so

$$\nabla_{\dot{\gamma}} J - \nabla_J \dot{\gamma} = [\dot{\gamma}, J] \tag{4.174}$$

Since it has just been shown that $[\dot{\gamma}, J] = 0$, one can write

$$\nabla_{\dot{\gamma}} J = \nabla_J \dot{\gamma} \tag{4.175}$$

Using this result, and the fact that $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ (since $\dot{\gamma}$ is a geodesic) one can write

$$\nabla_{\dot{\gamma}}^2 J = \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} J = \nabla_{\dot{\gamma}} \nabla_J \dot{\gamma} = \left[\nabla_{\dot{\gamma}}, \nabla_J\right] \dot{\gamma} \tag{4.176}$$

From this relationship, by using the definition of the curvature tensor (Eq. (4.155)) and, again, canceling the commutator $[\dot{\gamma}, J]$, one obtains

$$\nabla_{\dot{\gamma}}^2 J = R(\dot{\gamma}, J)\dot{\gamma} \tag{4.177}$$

This last relationship is the Jacobi equation, written in a compact manner.

It is worth mentioning about the normal component $J_\perp$ of J, (i.e. that component of J orthogonal to $\dot{\gamma}$ along the geodesic γ), that it is also a Jacobi vector field, because one can always write $J = J_\perp + \lambda \dot{\gamma}$. It follows immediately that the velocity $\dot{\gamma}$ satisfies the Jacobi equation, so that $J_\perp$ must satisfy the same equation. This often allows restricting the study only to normal Jacobi fields.

References

Badescu, V.: Control optimal in ingineria proceselor termice. Matrix Rom, Bucuresti (2003)

Beju, I., Soos, E., Teodorescu, P.P.: Tehnici de calcul vectorial cu aplicatii. Editura Tehnica, Bucuresti (1976)

Casetti, L., Pettini, M., Cohen, E.G.D.: Geometric approach to Hamiltonian dynamics and statistical mechanics. Phys. Rep. **337**, 237–341 (2000)

do Carmo, M.P.: Riemannian Geometry. Birkhauser, Boston (1993)

Ionescu-Pallas, N.: Relativitate generala si cosmologie, Ed. Stiintifica si Enciclopedica, Bucuresti (1980)

Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry. Wiley, New York (1991)

Landau, L.D., Lifshitz, E.M.: The Classical Theory of Fields, 4th edn. Butterworth-Heinemann, Oxford (2000)

Wald, R.M.: General Relativity. The University of Chicago Press, Chicago (1984)

Part II

Applications of Geometric Methods in Thermodynamics

Chapter 5
Equilibrium Thermodynamics

The foundations of the equilibrium thermodynamics have been established by Rudolf Clausius. In his understanding, the fundamental quantities are two extensive quantities, namely the internal energy U and the entropy S, while the main postulates are the first and the second laws of thermodynamics. There are many excellent introductory books in equilibrium thermodynamics, such as those of Landau and Lifshitz (1970) and Adkins (1997). Here, the principal results of the equilibrium thermodynamic are briefly presented, following the survey of Baylin (1994). Only those notions which are of interest for the next chapters are summarized.

5.1 Thermodynamic Potentials

In 1869, Francois Massieu proved that one single function is sufficient to describe consistently the entire edifice of equilibrium thermodynamics. From this function, through differentiation, one can get all thermodynamic relations. Following the terminology of mechanics, such a function was called *thermodynamic potential*. Massieu showed that the thermodynamic potential is not unique and pointed to two quantities which can play the role of potential:

- the Helmholtz free energy, F, given by the relationship:

$$F = U - TS \tag{5.1}$$

- the Gibbs free energy, G, defined by:

$$G = U - TS + PV \tag{5.2}$$

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_5

In these relationships, V is the volume of the thermodynamic system (*extensive quantity*) and P and T are the pressure and the temperature (*intensive quantities*).

To illustrate the method used by Massieu, the thermodynamic relationships will be deduced now by starting from Helmholtz free energy. The first law of thermodynamics for closed systems is written in the local form:

$$\delta Q - \delta W = dU \tag{5.3}$$

where δQ and δW represent very small amounts of heat and mechanical work transferred between the system and the environment and dU is the infinitesimal variation of the internal energy of the system. Consider the limit case, when the processes that occur in the system during the thermodynamic transformation are reversible. Then δQ and δW can be expressed by using state variables, through the relationships:

$$\delta Q = TdS, \quad \delta W = PdV \tag{5.4a, b}$$

By using relations (5.3) and (5.4) one can write again the expression of the first law of thermodynamics, under the form:

$$dU = TdS - PdV \tag{5.5}$$

The starting idea used by Massieu was the observation that the relationship (5.5) can be deduced by using purely mathematical arguments. So, the internal energy U can be considered from the very beginning as a function of two independent variables, namely V and S:

$$U = U(V, S) \tag{5.6}$$

Then, the pressure and the temperature can be defined by the relations

$$P \equiv -\left(\frac{\partial U}{\partial V}\right)_S, \quad T \equiv \left(\frac{\partial U}{\partial S}\right)_V \tag{5.7, 8}$$

Through differentiation of U and taking into consideration the relationships (5.6), (5.7) and (5.8) one obtains the relationship (5.5).

Massieu applied a similar procedure to the Helmholtz free energy F. By differentiating the relation (5.1) one obtains

$$dF = dU - TdS - SdT \tag{5.9}$$

By using the relation (5.5) one obtains:

$$dF = -SdT - PdV \tag{5.10}$$

By comparing the relationships (5.10) and (5.5) one sees that F may be considered as a function of two independent variables, T and V, i.e.:

$$F = F(T, V) \tag{5.11}$$

In this case, in order for the differential of F to be given by (5.10), the entropy and pressure must be defined by the relations:

$$S \equiv -\left(\frac{\partial F}{\partial T}\right)_V, \quad P = -\left(\frac{\partial F}{\partial V}\right)_T \tag{5.12, 13}$$

Relations (5.12) and (5.13) show how, by the differentiation of F, one can obtain two thermodynamic quantities.

The procedure may be repeated for other quantities. For example, the internal energy is obtained by replacing the entropy given by Eq. (5.12) in (5.1), i.e.

$$U = F - T\left(\frac{\partial F}{\partial T}\right)_V = -T^2\left[\frac{\partial (F/T)}{\partial T}\right]_V \tag{5.14}$$

The specific heat at constant volume, c_V, is obtained by using the standard definition, and followed by the usage of relation (5.4a). One finally obtains:

$$c_V \equiv \left(\frac{dQ}{dT}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_V \tag{5.15}$$

The idea of Massieu is important. Indeed, his procedure uses just one function, for example $F(T, V)$, from which all thermodynamic relations can be deduced. Instead, the initial procedure which was used by Clausius, has been based on two functions, for example the energy and the entropy.

The procedure proposed by Massieu can be used in case of the Gibbs free energy, G. Thus, from the relations (5.2) and (5.3) one finds the expression:

$$dG = -SdT + VdP \tag{5.16}$$

Relation (5.16) suggests that G may be considered as function of the independent variables T and P, i.e.:

$$G = G(T, P) \tag{5.17}$$

In this case, in order for the differential of G to be given by (5.16), the entropy and the volume must be defined by the equations:

$$S \equiv -\left(\frac{\partial G}{\partial T}\right)_P, \quad V = \left(\frac{\partial G}{\partial P}\right)_T \tag{5.18, 19}$$

The internal energy is obtained by replacing the entropy and the volume given by the relationships (5.18) and (5.19), respectively, in Eq. (5.2), i.e.

$$U = G - T\left(\frac{\partial G}{\partial T}\right)_P - P\left(\frac{\partial G}{\partial P}\right)_T \tag{5.20}$$

The specific heat at constant pressure, c_P, is obtained from the standard definition, followed by the successive usage of the relations (5.4a) and (5.18). Making all calculations leads to the following chain of relations:

$$c_P \equiv \left(\frac{dQ}{dT}\right)_P = T\left(\frac{\partial S}{\partial T}\right)_P = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_P \tag{5.21}$$

The manner of using the procedure proposed by Massieu allows emphasizing two important observations:

- By passing from the relationship (5.6) to the relation (5.11), and then to the relationship (5.17), the *independent* variables were changed from (S, V) to (T, V), and finally to (P, V).
- Changing the *dependent* variables, from U to F, and finally to G, was made in a systematic way, by using the following scheme. One starts from a generic potential, denoted $I(X, Y)$, of two independent variables, X and Y. For this potential, one can write $dI(X, Y) = AdX + BdY$, where A and B are partial derivatives. Next, one passes to a new potential, having the form $J \equiv I - AX$. Its differential is given by $dJ = dI - AdX - XdA$. Substituting in this expression the previous relation for dI, it is found that $dJ = -XdA + BdY$. From this, one can deduce that J must be function of A and Y, i.e. $J = J(A, Y)$.

It can be concluded that, by forming the new dependent variable $F = U - TS$, the independent variables automatically changed from (S, V) to (T, V). In mathematics, such a transformation, that change simultaneously both the dependent variable and the independent variable, is called *Legendre transformation*.

Note that not any transformation of thermodynamic coordinates is necessarily a Legendre transformation. Consider, for example, the case when the internal energy U is a function of T and V. This is comparable with the relationship (5.5), which gives the differential dU only if S is considered function of T and V. Then:

$$dU(T, V) = T\left(\frac{\partial S}{\partial T}\right)_V dT + \left[T\left(\frac{\partial S}{\partial V}\right)_T - P\right] dV \tag{5.22}$$

From Eq. (5.22), by the derivation of U, the following expressions are obtained:

$$\left(\frac{\partial U}{\partial T}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V, \quad \left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial S}{\partial V}\right)_T - P \tag{5.23, 24}$$

It is clear that the Eqs. (5.7) and (5.8) are simpler than the Eqs. (5.23) and (5.24). In addition, the acceptance of the Eq. (5.6), which stated the dependence $U = U(S, V)$, has two advantages:

- It follows directly from the first law of thermodynamics and
- Its differentiation makes possible to obtain the entire information about the thermodynamic system.

This is the reason why Josiah Willard Gibbs has considered the equation $U = U(S, V)$ as a particular form of the *fundamental thermodynamics relation*.

Note that the fundamental relationship $U = U(S, V)$ can be put, if certain mathematical conditions are met, under the form $S = S(U, V)$. Because of this, the theory can be developed in two equivalent directions, as follows:

- If the internal energy U is the dependent variable and the entropy S is the independent variable, it is said that the *energy representation* is used;
- if the relationship $S = S(U, V)$ is adopted, it is said that the *entropy representation* is used.

5.2 Open Systems. Chemical Potential

Massieu's ideas were continued by Gibbs who developed in 1870 the theory to the case of open systems. This development is briefly present in the following. Consider an open system, which contains particles of different species that can enter or exit the system. The number of particles of species a in the system is denoted N_a. In this case, the fundamental relationship will have the following form (in the energy representation):

$$U = U(S, V, N_1, N_2, \ldots) \tag{5.25}$$

The derivative of the internal energy in relation to the number of particles of species a is the *chemical potential* (or the *electrochemical potential*) of that species:

$$\mu_a \equiv \left(\frac{\partial U}{\partial N_a}\right)_{S,V,N_{b\neq a}} \tag{5.26}$$

where $N_{b\neq a}$ means that the numbers of particles in all species different from a are kept constant. With this notation, the differential of U becomes:

$$dU = TdS - PdV + \sum_a \mu_a dN_a \tag{5.27}$$

Similarly, after simple calculations it is found that the differential of the Helmholtz free energy takes the form:

$$dF = -SdT - PdV + \sum_a \mu_a dN_a \tag{5.28}$$

Differential relationships of the type (5.27) are called *Gibbs equations*. These relationships are not regarded as generalizations of the first law, but they refer to virtual transformations that can take place between equilibrium states. Relations (5.27) can be extended, by introducing the *generalized displacements* y_i. In this case, the fundamental relation of thermodynamics has the form (in energy representation):

$$U = U(S, V, N_a, y_1, y_2, \ldots) \tag{5.29}$$

and a very small amount of work dW transferred between the system and the environment during a reversible process can be expressed as:

$$dW = X_1 dy_1 + X_2 dy_2 + \cdots. \tag{5.30}$$

where X_i are the so-called *generalized forces*.

5.3 Fundamental Relations and the Euler Relationship

In summary, it can be said that in the approach initiated by Massieu, all the information about a closed thermodynamic system is contained within the functions (potentials) $F(T, V)$ and $G(P, T)$. In case of open thermodynamic systems, a full description requires using several additional variables, namely the particle numbers of different species, N_a, and the chemical potential of these species, μ_a. Therefore, if the number of species is n, in the description of the thermodynamic system are involved $2n + 5$ variables, i.e.:

$$P, T, V, U, S, N_a, \mu_a \quad (a = 1, \ldots, n) \tag{5.31}$$

Gibbs argued that among these variables, $n + 2$ are independent variables, which can be chosen arbitrarily. For example, the list of independent variables may consist of P, T and $N_a (a = 1, \ldots, n)$. The other $n + 3$ variables in (5.31) are considered dependent and their values can be obtained as function of the values of the independent variables, either by using theoretical relationships or by experimental procedures.

The choice of the independent variables is influenced by some practical reasons or by convenience. It should be noted that this choice dictates on the thermodynamic potential which should be used. For example, if the $n+2$ independent variables are S, V and $N_a (a = 1,\ldots,n)$, then the potential that can be used is:

$$U = U(S, V, N_a) \quad (a = 1,\ldots,n) \tag{5.32}$$

The other $n+3$ dependent variables are determined as follows. Relation (5.32) allows getting one of these dependent variables while the other $n+2$ dependent variables are obtained by differentiating the thermodynamic potential:

$$T = \left(\frac{\partial U}{\partial S}\right)_{V,N_a} \quad -P = \left(\frac{\partial U}{\partial V}\right)_{S,N_a}, \quad \mu_a = \left(\frac{\partial U}{\partial S}\right)_{S,V,N_{b\neq a}} (a = 1,\ldots,2) \tag{5.33}$$

The above considerations justify why Gibbs has designated the Eq. (5.32) as the fundamental relationship. The set of independent variables S,V and $N_a (a = 1,\ldots,n)$ is also called *set of natural coordinates*. On the other hand, it can be shown by direct calculation that not every set of $n+2$ independent variables has the properties of the natural coordinates. For example, if T, V and $N_a (a = 1,\ldots,n)$ are chosen as independent variables, the entropy S cannot be obtained by differentiation of $U(T, V, N_a)\,(a = 1,\ldots,n)$, but only $(\partial S/\partial V)_{T,N_a}$. Such sets of independent variables constitute *unnatural coordinates*.

The Eq. (5.32) is not the only fundamental relationship. A second fundamental relationship often used is $F = F(T, V, N_a)$. From this expression one can obtain the entropy S and the pressure P by using Eqs. (5.12) and (5.13), respectively, while the chemical potential μ_a is obtained from

$$\mu_a = \left(\frac{\partial F}{\partial N_a}\right)_{T,V,N_{b\neq a}} \tag{5.34}$$

Finally, the internal energy is obtained from the relationship $U = F + TS$.

Other fundamental relations often used are $G = G(P, T, N_a)$ and $H = H(S, P, N_a)$. A fifth fundamental relationship is the *Euler relationship*:

$$U - TS + PV - \sum_a \mu_a N_a = 0 \tag{5.35}$$

By the differentiation of (5.35) and using (5.27), the *Gibbs-Duhem relationship* is found:

$$-SdT + VdP - \sum_a N_a d\mu_a = 0 \tag{5.36}$$

5.4 Thermodynamic Stability

The theory of the stability of thermodynamic systems was developed by similarity with the theory of mechanical stability. In this regard, it is useful to recall some basic ideas on the mechanical equilibrium.

5.4.1 Mechanical Equilibrium

Consider, for simplicity, a discrete mechanical system consisting of a finite number of particles. The resultant of the forces acting on the particle i is denoted $\mathbf{F}_i$. If the force $\mathbf{F}_i$ is conservative, then there is a scalar potential U so that:

$$\mathbf{F}_i = -\nabla_i U \tag{5.37}$$

Assume that the particle i is initially in equilibrium. If the force $\mathbf{F}_i$ is null ($\mathbf{F}_i = 0$), then the particle will remain in equilibrium. From Eq. (5.37) one sees that $\nabla_i U = 0$, which means that U reaches its extreme value. Therefore, at that extreme, the variation in the function U is null, i.e.:

$$\delta U = 0 \tag{5.38}$$

The classic example is that of a mechanical system consisting of a single small, heavy ball located inside a spherical cavity. If U is identified in this case with the gravitational potential energy, the ball finds its equilibrium only when it is at the bottom of the spherical cavity.

5.4.2 Principles of Extreme in Thermodynamics

It was seen that in mechanics the notion of equilibrium is associated with the extreme of a potential. By similarity, the same thing happens in thermodynamics. Next, two examples are presented, which are often used in applications.

5.4.2.1 The Principle of Maximum Entropy

The Principle of Maximum Entropy was deduced by Gibbs starting from the Clausius inequality. This inequality says that the necessary and sufficient condition for an isolated thermodynamic system (possibly heterogeneous) to be in thermodynamic equilibrium, is that during all virtual variations in the system state that do not change its energy, the entropy variations are negative or null:

$$\delta S_{U,V,N_a} \leq 0 \tag{5.39}$$

Assume that a thermodynamic system is in equilibrium state. This state can be of three kinds, depending on the sign of $\delta S_{U,V,N_a}$, as follows:

- *Stable equilibrium*, if $\delta S_{U,V,N_a} < 0$ for all variations in the neighborhoods Z around the equilibrium state;
- *Neutral equilibrium*, if $\delta S_{U,V,N_a} = 0$ for certain variations in the neighborhoods Z around the equilibrium state and $\delta S_{U,V,N_a} < 0$ for the rest of the variations;
- *Unstable equilibrium* if $\delta S_{U,V,N_a} > 0$ for certain variations in the neighborhoods around the state of equilibrium.

5.4.2.2 The Principle of Minimum Internal Energy

The Principle of the Minimum Internal Energy is as follows: the necessary and sufficient condition for an isolated thermodynamic system (possibly heterogeneous) to reach the thermodynamic equilibrium, is that during all virtual variations in the system state, that do not change the system entropy, the energy variations are negative or null:

$$\delta U_{E,V,N_a} \geq 0 \tag{5.40}$$

In this case, the state of thermodynamic equilibrium of a system is of three kinds, depending on the sign of $\delta U_{S,V,N_a}$, as follows:

- Stable equilibrium, if $\delta U_{S,V,N_a} > 0$ for all variations in the neighborhoods Z around the equilibrium state;
- Neutral equilibrium, if $\delta U_{S,V,N_a} = 0$ for certain variations in the neighborhoods Z around the equilibrium state and $\delta U_{S,V,N_a} > 0$ for the rest of the variations;
- Unstable equilibrium if $\delta U_{S,V,N_a} < 0$ for certain variations in the neighborhoods around the state of equilibrium.

5.5 Non-equilibrium Quantities

Thermodynamic systems are found most often in non-equilibrium. The methods of equilibrium thermodynamics can be used in these situations, as long as the notion of equilibrium makes sense at least locally. This involves the decomposition of the system into subsystems which continue to have macroscopic nature, and between these subsystems the transfer of energy and mass is sufficiently low that they can be considered in first approximation as being in equilibrium. One says that the system obeys the local equilibrium hypothesis, if each of these subsystems can be considered in equilibrium.

All the *extensive quantities* can be used for the description of the local thermodynamic equilibrium states. Among them, $n+2$ are independent, for example U, V, N_a or S, V, N_a. By using these independent variables, some *intensive quantities* can be defined locally, such as the temperature, the pressure and the chemical potentials.

In addition, the description of the global non-equilibrium state requires the local definition of some *non-equilibrium quantities*, denoted $\alpha, \beta, \ldots$. The non-equilibrium quantities will be denoted collectively by α. Examples of such quantities are the deviations of the local temperature, or of the number of particles $N_{a\gamma}$ in the phase γ, relative to the average temperature values, or to the average number of particles, respectively, corresponding to the state of equilibrium. At thermodynamic equilibrium, all the non-equilibrium quantities become zero, i.e. $\alpha = 0$.

It is concluded that the local thermodynamic equilibrium can be described by functions such as $\mathrm{S}(U, V, N_a, \alpha)$ or $U(S, V, N_a, \alpha)$.

5.6 The Nature of the State of Thermodynamic Equilibrium

Consider a heterogeneous system in equilibrium. It may be conceptually split into two subsystems, denoted as first and second, respectively. Part of the first subsystem is conceptually transferred to the second subsystem. This will change the volume, the entropy and the particle number of each species, both in the first subsystem and in the second subsystem. The total system being in equilibrium, the variations of the volume, entropy and particle number of each species, at the level of the whole system, will be null, i.e.

$$\begin{aligned} \delta V &= \delta V' + \delta V'' = 0 \\ \delta S &= \delta S' + \delta S'' = 0 \\ \delta N_a &= \delta N_a' + \delta N_a'' = 0 \end{aligned} \tag{5.41}$$

The energy representation will be used. In this case, the following dependencies can be written:

$$U' = U'(S', V', N_a'), \quad U'' = U''(S'', V'', N_a'') \tag{5.42}$$

The variation of the internal energy of the total system, $\delta U = \delta U' + \delta U''$, will be determined next. After elementary calculations one finds:

$$\begin{aligned}\delta U = &\left(\frac{\partial U'}{\partial V'}\right)_0 \delta V' + \left(\frac{\partial U'}{\partial S'}\right)_0 \delta S' + \sum_a \left(\frac{\partial U'}{\partial N'_a}\right) \delta N'_a \\ &+ \left(\frac{\partial U''}{\partial V''}\right)_0 \delta V'' + \left(\frac{\partial U''}{\partial S''}\right)_0 \delta S' + \sum_a \left(\frac{\partial U''}{\partial N''_a}\right) \delta N''_a \end{aligned} \tag{5.43}$$

Here, the subscript 0 designates the initial state, before changing the configuration of the two subsystems. In case of the energy representation, the intensive variables are introduced in the usual way, as derivatives of the internal energy:

$$\left(\frac{\partial U'}{\partial V'}\right)_0 = -P', \quad \left(\frac{\partial U'}{\partial S'}\right)_0 = T', \quad \mu'_a = \left(\frac{\partial U'}{\partial N'_a}\right)_0 \tag{5.44}$$

Similar relationships occur for the quantities of the second subsystem. In addition, from the Eq. (5.41) it is deduced that $\delta V'' = -\delta V'$,$\delta S'' = -\delta S'$ and $\delta N''_a = -\delta N'_a$. Substituting these last results in Eq. (5.43), the variation of the internal energy δU of the entire system is obtained. For the system to be in stable thermodynamic equilibrium, the condition $\delta U \geq 0$ is needed, i.e.:

$$\delta U = (P'' - P')\delta V' + (T' - T'')\delta S' + \sum_a (\mu'_a - \mu''_a)\delta N_a \geq 0 \tag{5.45}$$

The equilibrium conditions and the stability conditions of the thermodynamic system will be analyzed one by one. To address the equilibrium conditions, note that relation (5.45) must be checked for any independent variation of $\delta V', \delta S', \delta N'_a$. Therefore, in order to obey the extreme condition, $\delta U = 0$, the following relations must be fulfilled:

$$P'' - P' = 0, \quad T' - T'' = 0, \quad \mu''_a - \mu'_a = 0 \tag{5.46}$$

The relations (5.46) are the equilibrium conditions for the thermodynamic system.

To address the stability conditions, first it is denoted by α or β any of the extensive quantities S, V or N_a. Then, one denotes by U_α the derivative of the internal energy in respect to the quantity α. It was previously shown that, in the new notation, $\delta\alpha'' = -\delta\alpha'$ (see relationships (5.41)). Consider, moreover, that $U'_{\alpha\beta} = U''_{\alpha\beta}$. Since at stable equilibrium the internal energy is a minimum, the stability condition is $\delta^2 U > 0$. After calculations, which are similar to those necessary for the deduction of relationship (5.45), one finds:

$$\delta^2 U = 2\left(\frac{1}{2}\right) \sum_\alpha \sum_\beta U'_{\alpha\beta} \delta\alpha' \delta\beta' > 0 \tag{5.47}$$

Here, the variations $\delta\alpha', \delta\beta'$ refer to $\delta V', \delta S', \delta N'_a$. Because these variations are arbitrary, they may be chosen as being null, except for only one of them. For different choices of the variations of the extensive quantities, it follows that, in order that the relationship (5.47) to be fulfilled, one needs that:

$$\frac{\partial^2 U}{\partial V^2} \geq 0, \quad \frac{\partial^2 U}{\partial S^2} \geq 0, \quad \frac{\partial^2 U}{\partial N_a^2} \geq 0 \tag{5.48a, b, c}$$

The relationships (5.48a, b, c) constitute the conditions for mechanical, thermal and chemical stability, respectively. Note that a single principle (i.e. the minimum internal energy principle) has been used to obtain all the three conditions of stability.

5.7 Other Extreme Principles. Availability Function

Stability of fluid phases can be addressed by using the concept of availability function. The principle of the minimum of the availability function is stated as follows. A necessary and sufficient condition for a homogeneous system to be in equilibrium in the state $E_0(U_0, S_0, V_0, N_{a0})$ is that the parameters T_0, R_0, μ_{a0} exist so that the availability function $A_0(Z)$, defined as follows:

$$A_0(Z) \equiv U - T_0 S + P_0 V - \sum_a \mu_{a0} N_a \tag{5.49}$$

is null for that state $Z = E_0$, and has non-negative variations of the first order for all neighboring states $Z(U, S, V, N_a, \alpha)$, whether they are states of equilibrium, or non-equilibrium states:

$$A_0(E_0) = 0, \quad \delta A_0 \geq 0 \tag{5.50a, b}$$

Here $A_0(E_0)$ is a notation for $A_0(U_0, S_0, V_0, N_{a0})$.

5.8 Another Form of the Stability Condition

The Euler equation (5.35) can be applied successively for two equilibrium states, $E(U, S, T, V, P, N_a)$ and $E_0(U_0, S_0, T_0, V_0, P_0, N_{a0})$, respectively. By subtracting the obtained expressions, the following result is found:

$$U - U_0 = TS - T_0 S_0 - PV + P_0 V_0 + \sum_a \mu_a N_a - \sum_a \mu_{a0} N_{a0} \tag{5.51}$$

Equation (5.51) is replaced in the *availability theorem* (5.50a, b), applied to the case of the equilibrium state E, yielding:

$$S(T - T_0) - V(P - P_0) + \sum_a N_a(\mu_a - \mu_{a0}) > 0 \tag{5.52}$$

The same reasoning can be applied to the equilibrium state E_0. Instead of (5.52) it is obtained:

$$S_0(T_0 - T) - V_0(P_0 - P) + \sum_a N_{a0}(\mu_{a0} - \mu_a) > 0 \tag{5.53}$$

Adding these equations and using the notations $\delta P_0 \equiv P_0 - P$, $\delta T_0 \equiv T_0 - T$, etc., it is found that:

$$\delta T \delta S - \delta P \delta S + \sum_a \delta\mu_a \delta N_a > 0 \tag{5.54}$$

Relation (5.54) is another form of the stability condition for the thermodynamic system. This result is due to Gibbs.

5.9 Applications for Systems in Contact

The principle of the minimum of the availability function can be used to study the time evolution of a thermodynamic system in contact with sources (or reservoirs) of different kinds. The thermodynamic system will be further denoted with K. Sources (reservoirs) are thermodynamic systems (usually considered to be very large), whose properties do not change over time, even if they are in contact with other thermodynamic systems. Common examples of sources are:

- (h) sources of heat, at temperature T_0;
- (v) sources of volume, at pressure P_0 and temperature T_0;
- (a) sources of particulate of species a, at potential chemical μ_{a0} and temperature T_0;
- (w) sources of work $\bar{W}_0$. This work will be referred to as "*special work.*" It is of a different kind than the *displacement work* (given, in case of reversible processes, by the relationship of the type PdV).

The classical cases considered by Gibbs refer to the sources of kind (h) and (v). In the following it is shown how the availability function can be used in this context. Any finite change of property will be denoted by Δ.

One can show that the properties U, S, V and N_a of the system K, being in contact with the above sources, evolve over time in such a way that they satisfy the following inequality (Baylin 1994):

$$\Delta U - T_0 \Delta S + P_0 \Delta V - \sum_a \mu_{a0} \Delta N_a - \bar{W}_0 = -\Delta Q' < 0 \tag{5.55}$$

In this relationship $\Delta Q'$ is a measure of the inequality and represents the lost work. Relation (5.55) can be expressed by using the availability function as follows:

$$\Delta A_0 - \bar{W}_0 \leq 0 \tag{5.56}$$

It is expected that the availability function decreases during the system evolution and cancels at equilibrium. In this later case, the system parameters will become equal with V_0, P_0 and μ_{a0}.

In order to demonstrate the relationship (5.55), assume that the total system (consisting of the system K and the sources with which it is in contact) is isolated. In this case, the variation of the total system volume and total number of particle species a is null, i.e.:

$$\Delta V + \Delta V_0 = 0, \quad \Delta N_a + \Delta N_{a0} = 0 \tag{5.57, 58}$$

Also, the total internal energy of the system is conserved:

$$\Delta U + \Delta U_h + \Delta U_v + \Delta U_w + \sum_a \Delta U_a = 0 \tag{5.59}$$

The entropy of the total system satisfies the second law of thermodynamics, which in mathematical terms has the form:

$$\Delta S + \Delta S_h + \Delta S_v + \Delta S_w + \sum_a \Delta S_a = T_0^{-1} \Delta Q' \geq 0 \tag{5.60}$$

The four sources satisfy the following Gibbs equations:

$$\begin{aligned} \Delta U_h - T_0 \Delta S_h &= 0 \\ \Delta U_v - T_0 \Delta S_v + P_0 \Delta V_0 &= 0 \\ \Delta U_a - T_0 \Delta S_a - \mu_{a0} \Delta N_{a0} &= 0 \\ \Delta U_w - T_0 \Delta S_w + \bar{W}_0 &= 0 \end{aligned} \tag{5.61a–d}$$

The calculation continues in the following way. Multiply the Eq. (5.60) by T_0 and the result is subtracted from the sum of the four relationships (5.61a–d). The result is put into a simpler form using (5.57) and (5.58). Finally, one gets the expected result, which is the relationship (5.55).

If, for the moment, the special work $\bar{W}_0$ is ignored, from (5.55) one can derive some relationships that characterize the time evolution of the systems:

$$
\begin{aligned}
(\Delta S)_{UVN_a} &\geq 0 \\
(\Delta U)_{SVN_a} &\leq 0 \\
(\Delta F)_{TVN_a} &\leq 0
\end{aligned}
\tag{5.62a–c}
$$

These relationships, which are consistent with the extreme principles already presented, highlight the importance of the thermodynamic potentials for the appropriate description of the trends of processes development.

5.10 Work Potentials

Gibbs was the first to notice that the functions U and F can be used as a measure of the ability of a thermodynamic system to do work, similarly with the potentials of the conservative forces in classical mechanics. The following comparison will make this clearer.

5.10.1 Work Potential in Mechanics

In mechanics, the work W done by a conservative force is linked with the variation of the potential energy U_{mec} of a system by means of the following relationship, which is a form of the energy conservation law:

$$W = -\Delta U_{mec} \tag{5.63}$$

Remember that, if the mechanical system consists of a single material point, the work is given by the curvilinear integral:

$$W = \int \mathbf{F} \cdot d\mathbf{r} \tag{5.64}$$

where $d\mathbf{r}$ is the element of trajectory of the material point and $\mathbf{F}$ is the force acting on it. If this force is conservative, one finds that:

$$\mathbf{F} = -\nabla U_{mec} \tag{5.65}$$

These simple relationships show that the potential energy U_{mec} plays the role of work potential.

5.10.2 Potentials of Special Work in Thermodynamics

The special work done on the system obeys the relationship (5.56), i.e. $\bar{W}_0 \geq \Delta A_0$. Therefore, ΔA_0 is the minimum amount of special work to be transferred from the environment towards a system that initially is in equilibrium, for changing the system properties with $\Delta U, \Delta S$, etc.

It is observed, however, that $\bar{W}_0$ initially appears in the relationship (5.61d), which refers to a reversible process. In that relationship, $\bar{W}_0$ may be replaced with $-W_0$, where W_0 is the work carried out *by* the system. In doing so, Eq. (5.56) becomes $W_0 \leq -\Delta A_0$, which means that the maximum special work that a system can perform on the environment, when the system exhibits the variations $\Delta U, \Delta S$, etc, is $-\Delta A_0$.

It is concluded that the thermodynamic evolution of the same system can be formulated both by a *maximum principle* and by a *minimum principle*:

$$\begin{aligned} W_0 &\leq -\Delta A_0 = -\Delta U + T_0 \Delta S - P_0 \Delta V + \sum_a \mu_{a0} \Delta N_a \\ \bar{W}_0 &\geq \Delta A_0 = \Delta U - T_0 \Delta S + P_0 \Delta V - \sum_a \mu_{a0} \Delta N_a \end{aligned} \tag{5.66, 67}$$

It is important to note that in thermodynamics the availability function A_0 is a measure of the systems ability to transfer work. Therefore, A_0 may be assimilated with a work potential. This potential occurs in relationships (5.66) and (5.67), which, from a mathematical point of view, are inequalities. In mechanics, the mechanical work potential U_{mec} appears within an equality (i.e. Eq. (5.63)).

In case that restrictions (constraints) exist, the general relations (5.66) and (5.67) take different forms, often encountered in practice. For example, when the entropy, the volume and the number of particles of each species are constant, the relationship (5.66) becomes:

$$W_0 \leq -(\Delta U)_{S,V,N_a} \tag{5.68}$$

If the temperature remains equal to the ambient temperature ($T = T_0$) and the volume and the number of particles of each species are constant, the relationship (5.66) becomes:

$$W_0 \leq -(\Delta F)_{T,V,N_a} \tag{5.69}$$

The right side of the relationships (5.68) and (5.69) cancels at equilibrium state, under the prescribed restrictions (constraints).

5.10.3 Total Work Potentials

The “special” work does not include the displacement work. The total work W transferred between the system and the environment can be defined as the sum of the displacement work and the special work:

$$W = \sum_i P_{0i} \Delta V_i + W_0 \tag{5.70}$$

The maximum principle can be reformulated for the total work, starting from the Eq. (5.66):

$$W \leq -\Delta U + T_0 \Delta S + \sum_a \mu_{a0} \Delta N_a \tag{5.71}$$

In the case that restrictions (constraints) exist, the general relationship (5.71) takes particular forms. For example, when the entropy and the particle number of each species are constant, the Eq. (5.71) becomes:

$$W \leq -(\Delta U)_{S,N_a} \tag{5.72}$$

If the temperature remains equal to the ambient temperature ($T = T_0$) and the number of particles of each species is constant, the Eq. (5.71) becomes:

$$W \leq -(\Delta F)_{T,N_a} \tag{5.73}$$

In equilibrium state, the right hand side of the relations (5.72) and (5.73) is canceled. In non-equilibrium state, that right side no longer cancels and relationships (5.72) and (5.73) can be used to define the thermodynamic forces:

$$X_\alpha(U) = -\left(\frac{\partial U}{\partial \alpha}\right)_{S,N}, \quad X_\alpha(U) = -\left(\frac{\partial F}{\partial \alpha}\right)_{T,N} \tag{5.74a, b}$$

Here α denotes a non-equilibrium variable.

5.11 The Principle of Maximum for Thermally Isolated Systems

It is assumed that a thermally isolated system, initially is in an unspecified state (0) and then evolves adiabatically towards a state of equilibrium characterized by the quantities S, V and N_a. During the evolution, the system can exchange work with the environment. For a thermally isolated system, the first law of thermodynamics is reduced to:

$$W = -\Delta U \tag{5.75}$$

Denote the final internal energy of the system by $U(S, V, N_a)$. The work performed by the system during the adiabatic transformation is

$$W = U_0 - U(S, V, N_a) \tag{5.76}$$

Consider different final states, having the same volume V and the same number of particles N_a, but having different values of the entropy S. Assume that in all cases the system starts from the same initial state. The different final states previously considered correspond to different relaxation processes. Using relation (5.76), one can calculate the work done by the system during these different relaxation processes.

From (5.76) one can see that:

$$\left(\frac{\partial W}{\partial S}\right)_{V,N} = -\frac{\partial U(S, V, N_a)}{\partial S} = -T < 0 \tag{5.77}$$

This relationship states that for the same volume V and the same number of particles N_a, the work W decreases if the final value of the entropy S increases. Therefore, to maximize the work transferred from the system towards the environment, the relaxation should tend towards a final state of minimum entropy. Since the entropy of an isolated system cannot decrease, it is concluded that the work can be maximized by choosing a process that ends with a value of the entropy equal to the initial value (if such a process is possible).

5.11.1 Connection with Gouy-Stodola Theorem

The previous result, concentrated in the relationship (5.77), is linked to the *Gouy-Stodola theorem*, which states that in an adiabatic process during which the entropy of a system increases with the value ΔS, the amount of available work of the environment decreases by the value $T_0 \Delta S$, where T_0 is the smallest available temperature.

References

Adkins, C.J.: Equilibrium Thermodynamics, 3rd edn. Cambridge University Press, Cambridge (1997)

Baylin, M.: A Survey of Thermodynamics. American Institute of Physics, New York (1994)

Landau, L.D., Lifshitz, E.M.: Statistical Physics, 2nd edn. Pergamon Press, Oxford (1970)

Chapter 6
Thermodynamics as a Contact Geometry Structure

6.1 Contact Manifolds

In the last twenty years, modern mathematical methods, previously employed successfully in other branches of physics, began to be used in the field of thermodynamics. Methods of differential geometry, differential forms and Lie groups have already gained some authority. They were used previously to systematize classical mechanics, relativity theory and dynamical systems theory.

The main methods that will be considered here are those that use the techniques of contact geometries (Mrugala 1978, 1984, 1993, 2000). Recently, Poisson geometries and Jacobi geometries have been also used.

The chapter closely follows the excellent presentation by Mrugala (2000). The fundamental mathematical concept that underlies the usage of any of these geometries in thermodynamics is the *thermodynamic phase space* (TPS). For a thermodynamic system with n degrees of freedom, the TPS has the dimension $2n+1$, being a manifold endowed with a contact structure as defined by a Gibbs 1-form θ, in which all parameters are independent.

The contact structure allows the introduction of a metric G over TPS. The thermodynamics significance of such metric is not clear so far. The metric G can be reduced for a Gibbs space of dimension $n+1$. This approach will not be discussed here. Also, the metric G can be reduced to a n-dimensional Legendre submanifold S of the TPS. From the physical point of view, S represents different states of several, possibly different, thermodynamic systems.

Several results of contact geometry are reminded now. They will be used later.

A $2n+1$-dimensional manifold is called *contact manifold* if it allows the existence of a differential 1-form θ with the property that:

$$\theta \wedge (d\theta)^n \neq 0 \tag{6.1}$$

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_6

where $\wedge$ represents the *exterior product* and $(d\theta)^n \equiv d\theta \wedge (n\ times) \wedge d\theta$. Condition (6.1) states that θ is non-degenerate. The form θ is called *contact form*.

According to the theorem of Darboux, there is a canonical (contact) local coordinate system (x^0, x^i, p_i), $i = 1, \cdots, n$, in which θ has the simplest canonical form:

$$\theta \equiv dx^0 + p_i dx^i \quad (i = 1, \cdots, n) \tag{6.2}$$

In this expression, and further, the Einstein summation convention is used.

The condition of non-degeneracy (6.1) is interpreted geometrically in many ways, but using Eq. (6.2) it is deduced that $\theta \wedge (d\theta)^n$ represents the volume form on the manifold M.

A 1- form θ defines on a manifold M a $2n$-dimensional distribution D, i.e. a field of tangent $2n$-dimensional hyperplanes D_m, so that the following relations can be written:

$$D = \underset{m \in M}{U} D_m,\ D_m = \{X \in T_m M;\ \ \theta(X) = 0\} \tag{6.3}$$

where X denotes a field of vectors on M and $T_m M$ is the tangent space to M in a point $m \in M$. Locally, D is given by $2n$ vector fields, for example:

$$P_k = \frac{\partial}{\partial p_k}, \quad \chi_k = \frac{\partial}{\partial x^k} - p_k \frac{\partial}{\partial x^0}, \quad (k = 1, \ldots, n) \tag{6.4}$$

The distribution D is called *contact distribution* or *contact structure* on the manifold M.

The contact structure D is not given by a unique 1-form θ. If ρ is a function that does not vanish on M, the 1-form $\rho\theta$ satisfies the degeneration condition (6.1) and defines the same field D of hyperplanes.

In thermodynamics, the most important role is played by the maximum dimensional integral submanifolds of the contact distribution D, which are called *Legendre submanifolds*, denoted further on by S. Their name comes from the fact that the Legendre transformations preserve the submanifolds S, that these transforms map any Legendre manifold onto itself. From the thermodynamic point of view, the dimension of any Legendre submanifold coincides with the number of thermodynamic degrees of freedom.

The next theorem states that from a geometric point of view, the condition (6.1) means that D is maximally non-integrable, i.e. the dimension of its integral submanifolds cannot exceed n.

Theorem 6.1 *Let (M, θ) be a contact manifold with $2n + 1$ dimensions. The maximal dimension of the integral submanifolds of the contact distribution D (or,* *equivalently, of the integral submanifolds of* $\theta = 0$*) is* n*. The proof of this theorem is found in Mrugala* (2000).

The Legendre submanifolds can be locally described by using a *generating function* Φ, according to the following theorem.

Theorem 6.2 *For any partition* $I \cup J$ *of the set of indices* $\{i, \ldots, n\}$ *in two disjoint subsets I and J, and for a function* $\Phi(p_I, \mathrm{X}^J)$ *of n variables* $p_i,\ i \in I$ *and* $x^j,\ j \in J$*, the* $n + 1$ *equations:*

$$x^i = \frac{\partial \Phi}{\partial p_i},\ p_{\mathrm{j}} = -\frac{\partial \Phi}{\partial x^j}\ x^0 = \Phi - p_i \frac{\partial \Phi}{\partial p_i} \tag{6.6}$$

define a Legendre submanifold S of M^{2n+1}*. Conversely, any Legendre submanifold of* (M, θ)*, in the neighborhood of any point, is defined by these equations, for at least one of the* 2^n *possible choices of the subset I. The proof of this theorem is found in Mrugala* (2000).

$\Phi(p_I, x^J)$ is called generating function for the Legendre submanifolds, because all the variables p_J, x^I and x^0 are completely defined by Φ and the derivatives of Φ. It is easy to see the connection between the solutions (6.6) and the Legendre transformations.

6.1.1 Formulation of the First Law of Thermodynamics

The first law of thermodynamics can be expressed in terms of the Legendre submanifolds as follows:

> Any thermodynamic system in equilibrium is represented in a TPS (M, θ) by the Legendre submanifolds of the equation $\theta = 0$.

It is obvious that a thermodynamic system is not represented by a single Legendre manifold. It will be represented by a n-dimensional surface composed of fragments of different Legendre submanifolds, one piece for each thermodynamics phase. Through every point in M there is an infinity of Legendre submanifolds. Only some of these submanifolds correspond to real thermodynamic systems. From Theorem 6.2 one sees that in the contact coordinates (see Eq. (6.2)), a given Legendre manifold S can be represented, in principle, in an equivalent manner, by 2^n functions Φ of n variables. These functions correspond to different thermodynamic potentials. Therefore, for given generating function Φ, the system of Eq. (6.6) can be interpreted as consisting of a fundamental relationship and n thermodynamic equations of state.

6.1.2 The Dual Approach

There is a single one-dimensional (1D) *characteristic distribution*, dual of D, denoted by Ξ, defined by a global *characteristic vector field* ξ, so that:

$$i_\xi d\theta = 0,\ i_\xi \theta = 1\ (\text{or } i_\xi(\theta \wedge (d\theta)^n) = (d\theta)^n) \tag{6.7}$$

where i_ξ is the *interior (internal, inner) product* (*contraction*) with ξ. In contact coordinates, the global characteristic vector field is:

$$\xi = \frac{\partial}{\partial x^0} \tag{6.8}$$

Thus, $TM = D \otimes \Xi$ where TM is the tangent bundle on M and D and Ξ are two complementary vector subbundles of TM. The fields (6.4) and (6.8) satisfy the following commutation relations:

$$\left[\chi_i, \chi_j\right] = \left[P_i, P_j\right] = [\chi_i, \xi] = [P_i, \xi] = 0,\ \left[\chi_\mathrm{i}, P_\mathrm{j}\right] = \delta_{ij}\xi \tag{6.9}$$

From the last relationship (6.9) one sees that D is not involutive.

6.2 Contact Transformations and Contact Vector Fields

Definition 6.1 A diffeomorphism $\lambda : M \to M$ is said to be a contact diffeomorphism if it preserves the contact distribution D of M, i.e. if λ checks

$$\lambda^* \theta = \rho\theta, \quad \lambda \in \Gamma \tag{6.10}$$

where ρ is a function which does not cancel on M, λ^* is the pull-back map induced by λ and Γ is a group of diffeomorphisms.

Note that $\lambda^*\theta$ is a contact form, because it is non-degenerate, i.e. $\rho\theta \wedge (d(\rho\theta))^n = \rho^{n+1}\theta \wedge (d\theta)^n \neq 0$. Therefore, λ preserves the contact structure, but does not preserve the contact form. The diffeomorphisms with $\rho = 1$ preserve the contact form, too, and are called *strict contact transformations.*

Similarly, by a one-parameter group of *continuous contact transformations* one means a subgroup of the maps $\lambda_t : M \to M$ of Γ, that preserves the contact distribution D, i.e.:

$$\lambda_t^* \theta = \rho_t \theta \tag{6.11}$$

where ρ_t is a function that does not cancel on M.

Define X as a generator of this one-parameter subgroup of Γ, that is X is given by the relationships:

$$(Xf)(m) = \frac{d}{dt}|_{t=0}\left[\lambda_t^* f(m)\right] = \frac{d}{dt}|_{t=0}[f(\lambda_t(m))],\ (\forall m \in M) \tag{6.12}$$

for any function f on M. Therefore X is a vector field associated with λ_t.

The definition (6.11) of λ_t can be expressed in terms of X in the following way:

$$L_X\theta \equiv \frac{d}{dt}|_{t=0}(\lambda_t^*\theta) = \tau_t\theta \tag{6.13}$$

where $\tau_t = d\rho_t/dt$ and L_X is a Lie derivative. What matters here is that $L_X\theta$ is a product of θ with a function τ_t on M (if τ_t is zero, it is said that the 1-form θ is invariant). This justifies the Definition 6.2.

Definition 6.2 A vector field X on M is called *contact vector field* if it preserves the contact structure D or, equivalently, if

$$L_X\theta = \tau\theta, \quad \text{adica} \quad L_X\theta \wedge \theta = 0 \tag{6.14}$$

One can prove that any contact vector field forms a Lie algebra and that it does not belong to the contact distribution.

Now, consider M as a one-dimensional principal fiber bundle, its fibers being the integral curves of ξ. Then θ becomes a connection form and D and Ξ are called *horizontal* and *vertical distributions*, respectively.

Any vector X can be decomposed into its horizontal and vertical components, hX and vX, respectively:

$$X = vX + hX, \quad \text{where} \quad vX \equiv \theta(X)\xi \quad \text{and} \quad hX = X - vX \tag{6.15}$$

This allows the introduction of covariant differentiation on M. For a function f of real variable on M, its covariant differential Df is defined concisely by the two following relationships that may be considered equivalent:

$$Df(X) = df(hX) \quad \text{sau} \quad Df = \mathrm{d}f - (\xi f)\theta \tag{6.16}$$

for any vector field X on M.

Definition 6.3 By *contact vector field* associated with a function f on M, it is understood a vector field X_f defined by:

$$i_{X_f}\theta \equiv \theta(X_f) = f, \quad i_{X_f}d\theta = -Df \tag{6.17}$$

These two equations define the horizontal component $hX_f \equiv \bar{X}_f$ and the vertical component vX_f, respectively, of X_f:

$$X_f = hX_f + vX_f = \overline{X}_f + f\xi \tag{6.18}$$

One can see that, in contact coordinates:

$$\begin{aligned} X_f &= \frac{\partial f}{\partial p_i}\frac{\partial}{\partial x^i} + \left(p_i\frac{\partial f}{\partial x^0} - \frac{\partial f}{\partial x^i}\right)\frac{\partial}{\partial p_i} + \left(f - p_i\frac{\partial f}{\partial p_i}\right)\frac{\partial}{\partial x^0} \\ \bar{X}_f &= \frac{\partial f}{\partial p_i}\frac{\partial}{\partial x^i} + \left(p_i\frac{\partial f}{\partial x^0} - \frac{\partial f}{\partial x^i}\right)\frac{\partial}{\partial p_i} - p_i\frac{\partial f}{\partial p_i}\frac{\partial}{\partial x^0} \end{aligned} \qquad (6.19, 20)$$

Note that X_f and $\bar{X}_f$ are combinations of the vector fields (6.4) and ξ:

$$\begin{aligned} X_f &= (P_i f)\chi_i - (\chi_i f)P_i + f\xi \\ \overline{X}_f &= (P_i f)\chi_i - (\chi_i f)P_i \end{aligned} \qquad (6.21)$$

From relations (6.21) it is clear that $\bar{X}_f$ belongs to the contact distribution while X_f does not belong. Note that the definitions (6.21) depend on the coordinates, while the definition (6.3) does not depend on coordinates.

From the definition of X_f and the property of the Lie derivative, $L_X = i_X d + d i_X$, it is seen that X_f is a contact vector field, because the following relationships are valid:

$$L_{X_f}\theta = d\, i_{X_f}\theta + i_{X_f}d\theta = df - Df = (\xi f)\theta \sim \tau\theta \qquad (6.22)$$

This shows that X_f is a generator of a continuous contact transformation on M with $\tau = \xi f$. In contrast, $\bar{X}_f$ is not a contact vector field, because

$$L_{\overline{X}_f}\theta = d\, i_{\overline{\mathrm{Xf}}}\theta + i_{\overline{\mathrm{Xf}}}d\theta = -Df = -df + (\xi f)\theta \neq \tau\theta \qquad (6.23)$$

Moreover, the field vectors X_f form a Lie algebra, because

$$\begin{aligned} L_{[X_f, X_g]}\theta &= \left[L_{X_f}, L_{X_g}\right]\theta = L_{X_f}((\xi g)\theta) - L_{X_g}((\xi f)\theta) \\ &= \left(L_{X_f}(\xi g) - L_{X_g}(\xi f)\right)\theta \sim \tau\theta \end{aligned} \qquad (6.24)$$

Instead, the fields $\bar{X}_f$ do not form a Lie algebra. Both fields, X_f and $\bar{X}_f$, can be used to define in thermodynamics the Poisson, Jacobi and other, weaker, structures.

In local coordinates, or in a coordinate independent manner, one can prove that X_f and $\bar{X}_f$ have the following important properties:

$$\begin{array}{lll} \text{(a)} & X_c = c\xi, \quad (X_1 = \xi); & \overline{X}_c = 0 \ (c = \text{constant}) \\ \text{(b)} & X_{-f} = -X_f; & \overline{X}_{-f} = -\overline{X}_f \\ \text{(c)} & X_{f+g} = X_f + X_g; & \overline{X}_{f+g} = \overline{X}_f + \overline{X}_g \\ \text{(d)} & X_{fg} = fX_g + gX_f - fg\xi; & \overline{X}_{fg} = f\overline{X}_g + g\overline{X}_f \\ \text{(e)} & X_f f = f(\xi f); & \overline{X}_f f = 0 \\ \text{(f)} & X_f f^n = nf^n(\xi f) & \end{array} \qquad (6.25)$$

Comparing X_f with the general form of a vector field X on M

$$X = \dot{x}^i \frac{\partial}{\partial x^i} + \dot{p}_i \frac{\partial}{\partial p_i} + \dot{x}^0 \frac{\partial}{\partial x^0} \tag{6.26}$$

one obtains the following $2n+1$ differential equations:

$$\begin{aligned} \dot{x}^i &= \frac{\partial f}{\partial p_i} \equiv P_i f \\ \dot{p}_i &= p_i \frac{\partial f}{\partial x^0} - \frac{\partial f}{\partial x^i} \equiv -\chi_i f \\ \dot{x}^0 &= f - p_i \frac{\partial f}{\partial p_i} \equiv f - p_i P_i f \end{aligned} \tag{6.27}$$

Therefore, the flow induced by X_f represents a kind of *contact Hamilton equation*s, having a *contact Hamiltonian f*. From (6.27) we see that, unlike the Hamiltonian flows in classical mechanics, the contact Hamiltonian flows depend not only on the derivatives of f, but also on f.

An important property of the contact Hamiltonian flows can be deduced from the property (6.25e) of X_f. This can be rewritten as $X_f f \equiv df\left(X_f\right) = f(\xi f)$, which means that in the general case X_f is not tangent to all surfaces of constant level $f =$ constant. X_f is tangent only to one constant level surface, namely to the surface on which $f = 0$.

Moreover, if it happens that a Legendre manifold S is contained by the surface of zero level of f, $S \subset f^{-1}(0)$, then X_f is tangent to S. In this case, the contact Hamilton equations can be interpreted as describing a thermodynamic process.

6.3 Brackets Structures in Thermodynamics

First, remind some notions of classical mechanics and classical *Poisson brackets*. In conservative Hamiltonian mechanics, the basic concept is represented by a $2n$-dimensional phase space P^{2n} endowed with a closed differential 2-form ω $(\omega^n \neq 0,\ d\omega = 0)$. The pair (P^{2n}, ω) is called *symplectic manifold* and ω is called *symplectic form*.

According to *Darboux theorem*, on P^{2n} there is a local canonical coordinate system (p_i, q^i) in which $\omega = dp_i \wedge dq^i$. The dynamics of a mechanical system with a given Hamilton function $H(p, q)$ is governed by a Hamiltonian vector field X_H, defined by H and ω according with the formula $i_{X_H}\omega = -dH$ where i is the interior product. In canonical coordinates X_H is given by:

$$X_H = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} \tag{6.28}$$

This means that the Hamilton equations of the system (the *Hamilton flow*) have the form:

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i} \tag{6.29}$$

One can prove that ω is invariant with respect to X_H, i.e. the *Lie derivative* of ω is zero:

$$L_{X_H}\omega = 0 \tag{6.30}$$

For any two smooth functions f and g on P it can be defined, in several equivalent ways, the following Poisson bracket $\{f, g\}$:

$$\{f, g\} = \omega(X_f, X_g) = X_f g = L_{X_f} g \tag{6.31}$$

The same Poisson brackets can be defined by using canonical coordinates, as follows:

$$\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} \tag{6.32}$$

Definition 6.4 A Poisson bracket on a smooth manifold N is a map $\{\ ,\ \}: C^\infty(N, R) \times C^\infty(N, R) \to C^\infty(N, R)$ on the space of real functions C^∞ on N, with the following properties (the dimension r of N is arbitrary):

(1) *bilinearity*, i.e. $\{f, \lambda g_1 + \mu g_2\} = \lambda\{f, g_1\} + \mu\{f, g_2\}$ takes place, whatever $\lambda, \mu \in R$;
(2) *antisimmetry* (*skew-symmetry*), i.e. $\{f, g\} = -\{g, f\}$;
(3) *Jacobi identity*, i.e. $\{f, \{g, h\}\} = \{g, \{h, f\}\} + \{h, \{h, g\}\} = 0$;
(4) *Leibniz rule*, i.e. $\{f, gh\} = g\{f, h\} + \{f, g\}h$.

It is said that $\{\ ,\ \}$ defines a Poisson structure on N, and that the pair $(N, \{\ ,\ \})$ is a Poisson manifold.

It can be shown, using the Definition 6.4, that in local coordinates $(y^1, \ldots, y^r)$ the Poisson bracket has the form:

$$\{f, g\} = \sum_{k,l=1}^{r} J^{kl}(y)\frac{\partial f}{\partial y^k}\frac{\partial g}{\partial y^l} = \sum_{k,l=1}^{r}\{y^k, y^l\}\frac{\partial f}{\partial y^k}\frac{\partial g}{\partial y^l} \tag{6.33}$$

where $J^{kl}(y)$ are certain functions on N, that check Definition 6.4. From (6.33) one sees that the Poisson bracket $\{f, g\}$ is known, if the Poisson brackets for the local coordinates are known. Moreover, f and g enters in the Poisson brackets only through their first order partial derivatives.

According with Darboux theorem, there is on N a canonical system of local coordinates $(p^1,\ldots,p^n,q^1,\ldots,q^n,z^1,\ldots,z^h)$, in which:

$$\{f,g\} = \sum_{k,l=1}^{n} \left(\frac{\partial f}{\partial p^k}\frac{\partial q}{\partial q^l} - \frac{\partial f}{\partial q^k}\frac{\partial g}{\partial p^l}\right) \quad (2n+h=r) \tag{6.34}$$

The Poisson manifolds form a larger class than that of the symplectic manifolds, since N is not required to be of even dimension. The Poisson brackets can be degenerated for any dimension of N. This can be seen from (6.33) and (6.34).

Definition 6.5 A Jacoby bracket on a differential manifold N is a map $\{\ ,\ \}: C^\infty(N,R) \times C^\infty(N,R) \to C^\infty(N,R)$ on the space of real functions C^∞ on N, having the first three properties of Poisson bracket but with the fourth property (the Leibniz rule) replaced by the weaker condition:

$$(4')\ \mathrm{Supp}\{f,g\} \subset \mathrm{Supp}\, f \cap \mathrm{Supp}\ g$$

Any Poisson structure on the contact manifold (M,θ) must be degenerated, since the dimension of M is $2n+1$. On (M,θ) one can define other algebraic structures, less symmetric than the Poisson brackets.

Definition 6.6 On a contact manifold (M,θ), the following brackets can be defined:

(a) *Jacobi brackets*:

$$\begin{aligned}\{f,g\} &= \theta([X_f,X_g]) = X_f(g) - \{\xi f\}g \\ &= -d\theta(X_f,X_g) + f(\xi g) - g(\xi f)\end{aligned} \tag{6.35}$$

(b) *Cartan brackets*

$$[f,g] = \bar{X}_f(g) = -d\theta(X_f,X_g) = -d\theta(\bar{X}_f,\bar{X}_g) \tag{6.36}$$

(c) *Lagrange brackets*

$$(f,g) = \bar{X}_f(g) = \{f,g\} + g\{\xi f\} = [f,g] + f(\xi g) \tag{6.37}$$

These brackets can be defined in terms of the vector fields P_i, χ_i and ξ as follows:

$$\begin{aligned}\{f,g\} &= (P_i f)\chi_i g - (\chi_i f)P_i g + f\xi g - g\xi f; \\ [f,g] &= (P_i f)\chi_i g - (\chi_i f)P_i g \\ (f,g) &= (P_i f)\chi_i g - (\chi_i f)P_i g + f\xi g\end{aligned} \tag{6.38}$$

The latter relationships are useful because they highlight the symmetry properties of the brackets. The three brackets can be defined in local coordinates, too:

$$\begin{aligned}\{f,g\} &= \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i} - \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} + p_i\left(\frac{\partial f}{\partial x^0}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^0}\right)\\ &\quad + f\frac{\partial g}{\partial x^0} - g\frac{\partial f}{\partial x^0}\\ [f,g] &= \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i} - \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} + p_i\left(\frac{\partial f}{\partial x^0}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^0}\right)\\ (f,g) &= \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i} - \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} + p_i\left(\frac{\partial f}{\partial x^0}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^0}\right) + f\frac{\partial g}{\partial x^0}\end{aligned} \tag{6.39–41}$$

Remark 6.1 In the subspace $C_0^\infty(M,R) \subset C^\infty(M,R)$ of the first integrals of ξ, i.e. for f and g so that the following relations take place:

$$\xi f \equiv L_\xi f = \xi g \equiv L_\xi g = 0, \tag{6.42}$$

all these brackets are reduced to the standard Poisson brackets

$$\{f,g\}_{d\theta} = -d\theta(X_f, X_g) \tag{6.43}$$

This can be demonstrated with the help of the local coordinates representation:

$$\{f,g\} = [f,g] = (f,g) \equiv \{f,g\}_{d\theta} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i} - \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} \tag{6.44}$$

From the local expressions of these brackets it is observed that:

- { , } has all the properties of the Poisson bracket, but the Leibniz rule;
- [,] does not obey the Jacobi identity;
- (,) is not antisymmetric and does not obey the Jacobi identity;
- (,) is bilinear and obey the Leibniz rule, but only in the second entry;
- only (,) produces directly a flow contact Hamiltonian because the components X_f are equal to:

$$\dot{x}^i = (f, x^i), \quad \dot{p}_i = (f, p_i), \quad \dot{x}^0 = (f, x^0) \tag{6.45}$$

The general relations presented in this section will be used in applications in the next section.

6.4 Examples of Contact Flows in Thermodynamics

In thermodynamics, the manifold M of the previous section is usually a subset of R^{2n+1}. In the energy representation (in which internal energy $U(S, V)$ is a function of entropy and volume), the following correspondence between the variables of Sect. 6.3 and the thermodynamic quantities exists:

$$\begin{aligned}&(x^0;\ x^1,\ldots,x^n;\ p_1,\ldots,p_n)\\ &\Leftrightarrow (U;\ S, V, N_1,\ldots,N_{n-2};\ -T, P, -\mu_1,\ldots,-\mu_{n-2})\end{aligned} \tag{6.46}$$

Also, the differential 1-form θ is given by:

$$\theta = dU - TdS + PdV - \mu_k dN^k \quad (k = 1,\ldots,n-2) \tag{6.47}$$

The $2n + 1$ variables are assumed independent, except the case when restriction to a Legendre submanifold is made.

The entropy representation $S(U, V)$ can be also used, the correspondence between the variables in Sect. 6.3 and thermodynamic quantities being easy to establish.

The obvious applications of X_f in thermodynamics are:

- X_f induces a continuous contact transformation of M, i.e. a map of the manifold M on itself, so $L_{X_f}\theta = (\xi f)\theta \sim \theta$;
- As mentioned, X_f or the bracket (,) allows to find the thermodynamic contact Hamiltonian equations $\dot{x}^i = (f, x^i)$, $\dot{p}_i = (f, p_i)$ and $\dot{x}_0 = (f, x_0)$;
- $(f, g) = X_f g = L_{X_f} g$ allows finding first integrals of the flow induced by X_f. If, $(f, g) = 0$ and $(f, h) = 0$, then $(f, gh) = 0$.

For applications, the following theorem is important, whose demonstration can be found in Mrugala (2000).

Theorem 6.3 *Assume S is a Legendre submanifold of a contact manifold* (M, θ). *X_f is the tangent to S if and only iff cancels on S, i.e.* $S \subset f^{-1}(0)$.

Further examples will be given about the physical meaning of X_f. Also, the associated contact Hamilton equations will be built. A simple thermodynamic system is considered, characterized by usual thermodynamic quantities and, in addition, by the constant R (which is the perfect gas constant).

Example 6.1 For $f = U - TS + RNT - \mu N$, by using (6.19), (6.45) and (6.47) one finds:

$$X_f = (S - RN)\frac{\partial}{\partial S} + N\frac{\partial}{\partial N} + P\frac{\partial}{\partial P} + RT\frac{\partial}{\partial \mu} + U\frac{\partial}{\partial U} \tag{6.48}$$

and the contact Hamilton Eq. (6.27) have the form:

$$\dot{T} = \dot{V} = 0, \quad \dot{P} = P, \quad \dot{\mu} = RT, \quad \dot{S} = S - RN, \quad \dot{N} = N, \quad \dot{U} = U \tag{6.49}$$

Their integral curves are given by:

$$\begin{aligned} &T = T_0, \quad P = P_0 e^t, \quad \mu = RT_0 t + \mu_0 \\ &S = (S_0 - RN_0 t)e^t, \quad V = V_0, \quad N = N_0 e^t, \quad U = U_0 e^t \end{aligned} \tag{6.50}$$

For an ideal gas $f = 0$. Then, X_f is tangent to the Legendre submanifold S which represents this ideal gas and describes a "thermodynamic process" at constant volume V_0 and constant temperature T_0. It is easy to verify that this "process" preserves all thermodynamic relationships for ideal gases, such as:

$$PV = NRT, \quad U = \frac{3}{2}NRT, \quad U = TS - PV + \mu N \tag{6.51}$$

Example 6.2 For $f = NRT - (2/5)TS - (2/5)\mu N$, one finds:

$$X_f = \left(\frac{2}{5}S - RN\right)\frac{\partial}{\partial S} + \frac{2}{5}N\frac{\partial}{\partial N} - \frac{2}{5}T\frac{\partial}{\partial T} + \left(RT - \frac{2}{5}\mu\right)\frac{\partial}{\partial \mu} \tag{6.52}$$

The integral curves of X_f take the form:

$$\begin{aligned} &S = (S_0 - RN_0 t)e^{2t/5}, \quad V = V_0, \quad N = N_0 e^{2t/5}, \\ &T = T_0 e^{-2t/5}, \ P = P_0, \quad \mu = (\mu_0 + RT_0 t)e^{-2t/5}, \quad U = U_0 \end{aligned} \tag{6.53}$$

These relationships describe a "process" which is isobaric, isochoric and isoenergetic. Note that relations (6.51) keep their validity.

In Examples 6.1 and 6.2, the functions f were chosen so that the Legendre submanifolds S of the perfect gas are placed on the level hypersurfaces $f^{-1}(0)$. Consequently, X_f is tangent to S and could be treated as a "thermodynamic process". The situation is different if S is not placed on $f^{-1}(0)$. In the following examples X_f is not tangent to S and therefore cannot be treated as a generator of a thermodynamic process. Rather, X_f can be considered as a generator for a one-parameter family of thermodynamic systems.

Example 6.3 Assume that f is an affine function of the intensive parameters: $f = a + b^i p_i$. Then:

$$\dot{x}^i = b^i, \quad \dot{p}_i = 0, \quad \dot{x}_0 = a \tag{6.54}$$

and, as a consequence:

$$x^i = x^i_0 + b^i t, \quad p_i = p_{i0}, \quad x^0 = x^0_0 + at \tag{6.55}$$

Therefore, the intensive parameters are kept constant while the extensive parameters are linear functions of t. None of the Eq. (6.51) is kept in this case. Instead, X_f generates a continuous one-parameter family of thermodynamic systems (or, in other words, a one-parameter family of Legendre submanifolds S_t). A remarkable situation appears for $f = bP$. Then $V = V_0 + bt$ and all other parameters are fixed. For a fixed value of b, S_t represents a one-parameter family of hard spheres gases.

Example 6.4 If $f = a + b_i x^i$ is an affine function of extensive parameter, then X_f belongs to a class of contact vector fields. The integral curves of X_f take the form:

$$\dot{x}^i = x^i_0, \quad \dot{p}_i = p_{i0} - b_i t, \quad \dot{x}^0 = x^0_0 + (a + b_i x^i_0)t \tag{6.56}$$

and they do not represent a thermodynamic process. In this case, the physical meaning of X_f is not clear.

Example 6.5 Assume that $f = x^0 - \phi(x^1, \cdots, x^n)$. Then, the integral curves of X_f are:

$$\dot{x}^i = 0, \quad \dot{p}_i = p_i + \frac{\partial \phi}{\partial x^i}, \quad \dot{x}^0 = x^0 - \phi \tag{6.57}$$

In this case X_f produces a one-parameter family of Legendre submanifolds S_t, for a given manifold S. Assume that $\phi(x^1, \ldots, x^n)$ is such that $x^0 = \phi(x^1, \cdots, x^n)$ represents the fundamental relation of the thermodynamic system. Then $X_f\big|_S = 0$ and, of course, S is preserved.

Example 6.6 Assume that $f_1 = bP$, with b a non-negative constant; then the integral curves of $X_{f_1} = b\partial/\partial V$ are so that all parameters are preserved, except the volume V, that changes according with the relationship $V = V_0 + bt$. Therefore, X_{f_1} maps an ideal gas into a gas of hard spheres without interactions.

If $f_2 = -aV^{-1}$ $(a > 0)$, the integral curves of the field

$$X_{f_2} = -\frac{a}{V}\frac{\partial}{\partial U} - \frac{a}{V^2}\frac{\partial}{\partial P} \tag{6.58}$$

are so that (the new parameter τ is to be noticed):

$$U = U_0 - \frac{a}{V_0}\tau, \quad P = P_0 - \frac{a}{V_0^2}\tau \tag{6.59}$$

while all other parameters are preserved. It can be said that X_{f_2} maps a perfect gas into a gas of point-like particles that interact each other.

If one takes $f = f_1 + f_2 = bP - aV^{-1}$, the integral curves of X_f are so that T, S, N and μ do not change, while the following relationships occur

$$V = V_0 + bt, \quad U = U_0 - \frac{a}{b} \ln \frac{V_0 + bt}{V_0}, \quad P = P_0 - \frac{at}{V_0(V_0 + bt)} \tag{6.60}$$

The equation of state of the perfect gas ($P_0 V_0 = N_0 R T_0$) is no longer conserved. It turns into:

$$(P + \frac{at}{V(V - bt)})(V - bt) = NRT \tag{6.61}$$

which, for $t = 1$ resembles the Van der Waals equation of state. For fixed a and b one obtains a family of one-parameter Van der Waals gases.

Example 6.7 Two other forms of the Van der Waals gas are obtained if, instead of a transformation induced by $X_{f_1 + f_2}$, two consecutive transformations are assumed: that of X_{f_1}, followed by that of X_{f_2}, and vice versa. Two different two-parameter transformations are obtained, since the transformations induced by f_1 and f_2 do not commute. This can be seen from the Lie bracket:

$$\left[X_{f_1}, X_{f_2}\right] = \left[b\frac{\partial}{\partial V}, -\frac{a}{V}\frac{\partial}{\partial U} - \frac{a}{V^2}\frac{\partial}{\partial P}\right] = \frac{ab}{V^2}\frac{\partial}{\partial U} + \frac{2ab}{V^3}\frac{\partial}{\partial P} \neq 0 \tag{6.62}$$

or from the Jacobi bracket (6.39)

$$\{f_1, f_2\} = \left\{bP, -\frac{a}{V}\right\} = \frac{\partial f_1}{\partial P}\frac{\partial f_2}{\partial V} = \frac{ab}{V^2} \neq 0 \tag{6.63}$$

In case X_{f_1} is followed by X_{f_2}, instead of (6.61) a two-parameter family of equations of state is obtained:

$$\left(P + \frac{a}{V^2}\tau\right)(V - bt) = NRT \tag{6.64}$$

A different result is obtained if X_{f_2} is followed by X_{f_1}:

$$\left(P + \frac{a}{(V - bt)^2}\tau\right)(V - bt) = NRT \tag{6.65}$$

Equation (6.64) reproduces exactly the Van der Waals equation of state.

References

Mrugala, R.: Geometrical formulation of equilibrium phenomenological Thermodynamics. Rep. Math. Phys. **4**, 419 (1978)

Mrugala, R.: On the equivalence of two metrics in classical thermodynamics. Phys. A **125**, 631 (1984)

Mrugala, R.: Continuous contact transformations in thermodynamics. Rep. Math. Phys. **33**(1), 149–154 (1993)

Mrugala, R.: Geometrical methods in thermodynamics. In: Sieniutycz, S., De Vos, A. (eds.) Thermodynamics of energy conversion and transport, pp. 257–285. Springer, New York (2000)

Chapter 7
Thermodynamic Distance

The distance between two thermodynamic states can be defined in several ways. An approach that uses statistical tools is presented in the following (Diósi and Salamon 2000).

7.1 Classical Theory of Statistical Distance

Consider discrete classical statistical ensembles. They are parameterized by the normalized probability distributions $\mathbf{p} = (p_1, \ldots, p_k, \ldots)$. The space of parameters is the hyperplane:

$$\sum_k p_k = 1, \qquad (p_k \geq 0) \tag{7.1}$$

An alternative parametrization $\mathbf{c} = (c_1, \ldots, c_k, \ldots)$ can be introduced, where the components of the vector $\mathbf{c}$ are the square roots of the probabilities $c_k = p_k^{1/2}$ for $k = 1, 2, \ldots$. Then, the space of parameters (7.1) becomes spherical:

$$||\mathbf{c}|| = 1, \qquad c_k \geq 0 \tag{7.2}$$

i.e. by varying the components of the vector $\mathbf{c}$, a sector is generated on the surface of the unit hypersphere.

In practice, the probabilities p_k represent the relative frequencies x_k of some events of a given sample. The statistical distinguishability of the ensembles requires that the probabilities (7.1) are treated as relative frequencies in a larger sample of dimension N.

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_7

7.2 Parameterized Statistics

Assume that the statistical ensembles of probability distribution $\mathbf{p}$ are parameterized by a finite number of parameters $\mathbf{y} = (y^1, \ldots, y^n)$. These distributions represent a n-dimensional submanifold on the hypersurface of the sphere (7.2) and this hypersurface does not inherit a Euclidean geometry but a Riemannian geometry. The metric and the curvature that can be attached to this Riemannian geometry depend on $\mathbf{y}$.

The distance between two elements, $\mathbf{p}(\mathbf{y}_i)$ and $\mathbf{p}(\mathbf{y}_f)$, must be measured along a path that entirely belongs to the submanifold whose distribution is described by $\mathbf{y}$. The distance dl between infinitely close elements remains the same. If the elements are $\mathbf{y}$ and $\mathbf{y} + d\mathbf{y}$, it can be shown that:

$$dl = 2N^{1/2}\|dc\| \tag{7.3}$$

where

$$dc = \sum_{k=1}^{n} \frac{\partial c}{\partial y^k} dy^k \tag{7.4}$$

In the space of the parameters $\mathbf{y}$, the following Riemannian metric, $g_{ik}(\mathbf{y})$, can be defined:

$$g_{ik} = 4N \sum_r \frac{\partial c_r}{\partial y^i} \frac{\partial c_r}{\partial y^k} \tag{7.5}$$

In this case, the infinitesimal statistical distance (7.3) can be written in standard Riemannian form:

$$dl = \left(\sum_{i=1}^{n} \sum_{k=1}^{n} g_{ik}(\mathbf{y}) dy^i dy^k \right)^{1/2} \tag{7.6}$$

These equations describe a Riemannian geometry on the manifold of the parameterized discrete distributions $\mathbf{p}(\mathbf{y})$. The expressions can be rewritten by using probability distributions. In this case, the statistical metric (7.5) is:

$$g_{ik}(y) = 4N \sum_r \frac{\partial c_r}{\partial y^i} \frac{\partial c_r}{\partial y^k} = N \sum_r p_r \frac{\partial \ln p_r}{\partial y^i} \frac{\partial \ln p_r}{\partial y^k} \tag{7.7}$$

The compound forms of this expression are:

$$g_{ik}(\mathbf{y}) = N < \frac{\partial \ln p(\mathbf{y})}{\partial y^i} \frac{\partial \ln p(\mathbf{y})}{\partial y^k} > \; = -N < \frac{\partial^2 \ln p(\mathbf{y})}{\partial y^i \partial y^h} > \tag{7.8}$$

where $< \, >$ represent expectation values, calculated by using the probability distribution $\mathbf{p}(\mathbf{y})$.

The earlier expressions have applicability in the case of continuous distributions. Let $p(\mathbf{\Gamma}; \mathbf{y})$ be the probability distribution of the continuous random variables $\mathbf{\Gamma}$, depending on the continuous parameters $\mathbf{y}$. The normalization condition for the probability is:

$$\int p(\mathbf{\Gamma}; \mathbf{y}) d\mathbf{\Gamma} = 1 \tag{7.9}$$

Equation (7.8) will have the following explicit form:

$$g_{ik}(\mathbf{y}) = -N \int \frac{\partial^2 \ln p(\mathbf{\Gamma}; \mathbf{y})}{\partial y^i \partial y^k} p(\mathbf{\Gamma}; \mathbf{y}) d\mathbf{\Gamma} \tag{7.10}$$

7.3 Gibbs Statistics and Thermodynamics

Phenomenological thermodynamics can be derived from the statistical physics of particle systems. Consider a macroscopic system containing M moles of substance. If the system is in equilibrium, then its state $\mathbf{\Gamma}$ in the phase space is characterized by the probability distribution:

$$p(\mathbf{\Gamma}, \mathbf{y}) = \exp[-\Phi(\mathbf{y}) - \mathbf{y}\mathbf{F}(\mathbf{\Gamma})] \tag{7.11}$$

where $\mathbf{y} = (y^1, \ldots, y^n)$ are intensive parameters and $\mathbf{F}(\mathbf{\Gamma}) = (F_1(\mathbf{\Gamma}), \ldots, F_n(\mathbf{\Gamma}))$ are conjugated quantities that are conserved during the system motion. The function $\Phi(\mathbf{y})$ ensures the fulfillment of the normalization condition (7.9).

When the quantity $M \to \infty$, the ratio $\Phi(\mathbf{y})/M$ must converge towards the phenomenological thermodynamic potential $\varphi(\mathbf{y})$, per mole of system:

$$\frac{\Phi(\mathbf{y})}{M} \underset{M \to \infty}{\longrightarrow} \varphi(\mathbf{y}) \tag{7.12}$$

This is the so-called *thermodynamic limit*, which is the essence of statistical theory of Gibbs.

Compute now the statistical matrix of the Gibbs ensembles. Substituting (7.11) in (7.8), one finds:

$$g_{ik}(\mathbf{y}) = N\frac{\partial^2\Phi(\mathbf{y})}{\partial y^i \partial y^k} \tag{7.13}$$

If the system is large enough, the thermodynamic limit is valid and Φ can be replaced in (7.13) by $M\varphi$. This makes appearing a factor MN in the equation. Next, M is "absorbed" in N to yield:

$$g_{ik}(\mathbf{y}) = N\frac{\partial^2\varphi(\mathbf{y})}{\partial y^i \partial y^k} \tag{7.14}$$

This metric determines the statistical distance between different Gibbs ensembles, for different parameters $\mathbf{y}$. It is expressed in terms of the second derivatives of the thermodynamic potential.

An important aspect should be emphasized: the concept of statistical distance has induced the concept of distance between thermodynamic states. This last distance was previously used by Weinhold (1975) and Ruppeiner (1995).

It is instructive to get the Riemann metric by starting from the extensive parameters $x_k = -\partial\varphi/\partial y^k$ $(k = 1, \ldots, n)$, instead of starting from the intensive parameters. The convention is that the intensive parameters have high indices (superscripts) and the extensive parameters have low indices (subscripts). Therefore the metric tensor will have low indices for the intensive parameters and high indices for the extensive parameters. The result is:

$$g^{ik}(\mathbf{x}) = -N\frac{\partial^2 s(\mathbf{x})}{\partial x_i \partial x_k} \tag{7.15}$$

Here $s(\mathbf{x})$ is the function of specific entropy, related to the specific thermodynamic potential $\varphi(\mathbf{y})$ by the Legendre transformation:

$$s = \mathbf{x} \cdot \mathbf{y} - \varphi \tag{7.16}$$

The metric (7.15) can be called macroscopic, since it consists of second derivatives of the macroscopic entropy, with respect to the extensive variables of the macroscopic system. Note that the matrix of the metric contains second order derivatives, in both (7.14) and (7.15). This is true only for the entropy and its complete Legendre transformation, not for any partial Legendre transformations.

Next, briefly remind some terminological conventions. An extensive specific quantity is obtained by dividing the corresponding extensive quantity by the number of moles. Extensive density (in short, density), is called the corresponding extensive quantity divided by the volume. For simple systems in equilibrium, the number of independent extensive parameters exceeds by one the number of the specific extensive quantities.

It is useful to have a convenient expression for the metric, in terms of the extensive macroscopic variables $X_k = Nx_k$ $(k = 1, \ldots, n)$ and $X_{n+1} = N$, instead of the specific extensive parameters $s, x_1, \ldots x_n$. From (7.15) it is obtained:

$$g^{ik}(\mathbf{X}) = -\frac{\partial^2 S(\mathbf{X})}{\partial X_i \partial X_k} \tag{7.17}$$

where $S(\mathbf{X}) = Ns(\mathbf{x})$ is the extensive entropy function. Because:

$$\frac{\partial^2 S}{\partial^2 X_{n+1}^2} = \frac{\partial^2 S}{\partial N^2} = 0 \tag{7.18}$$

one finds that $g^{ik}(\mathbf{X})$ is degenerate, i.e. there are directions along which the metric measures a null distance. These directions always correspond to scaling one phase of the system. If the scale of the system is not fixed, the structure is only semi-Riemannian. Null directions result from the linear growth of the entropy, when any of the phases is scaled. This linearity makes the second derivatives, and finally the appropriate components of the metric tensor, to vanish along those directions.

Weinhold (1975) noted that by forming a suitable combination of such null directions, by simultaneously scaling two phases of a pure substance, a phase transition may be represented.

7.4 The Relevance of Riemann Geometry in Thermodynamics

Weinhold (1975) introduced the following metric for the phenomenological thermodynamics:

$$G^{ik}(\mathbf{Z}) = \frac{\partial^2 U(\mathbf{Z})}{\partial Z_i \partial Z_k} \tag{7.19}$$

where U is the internal energy and $\mathbf{Z} = (S, V, N_1, N_2, \cdots) \in R^{n+1}$ is the vector of the extensive variables of the system in the energy representation. Weinhold used the geometry only locally, to express in a convenient way the relations between the variations of the state variables. Later, Salamon and coworkers observed that the distances calculated for the perfect gas by using the metric G are actually familiar expressions of the variations of the kinetic energy of the gas molecules as a result of the passing of a shock wave. This then led to the idea of a connection between geometry and dissipation (Diósi and Salamon 2000).

Ruppeiner (1979) used the metric g of (7.15) to extend the scale for which the thermodynamic fluctuations theory can be used.

The metrics g from (7.15) and G from (7.14), respectively, are conformally equivalent (this means that the squares of the length elements differ by a scale factor—possibly dependent on position). However, it turns out that the metric g is more fundamental.

Similarly to the role of the energy representation in phenomenological thermodynamics, G serves mainly as an useful tool in computing $g(\mathbf{X})$, where $\mathbf{X} = (U, V, N_1, N_2, \cdots) \in R^{n+1}$ is the vector of the extensive variables in the entropy representation, for which $S = S(\mathbf{X})$ is the complete information.

Some issues that involve the notion of thermodynamics distance will be presented next, such as:

- fluctuation theory (which will be developed in the subsequent chapters);
- interpretation of dissipation;
- discussion of several horse-carrot-like theorems, which connect the dissipation in a system with imposing some conditions concerning the passage of the system through a given sequence of states.

7.5 A Covariant Theory of Fluctuations

In this section some elements of the theory of fluctuations are presented, but only when they are related with the notion of thermodynamic distance. The fluctuation theory will be addressed in more detail in the following chapters.

Consider the traditional expression for a fluctuation in a subsystem of volume V placed in a larger system of volume $V_0 \sim \infty$ and extensive densities $\mathbf{x}_0$. The probability that the subsystem have the extensive densities $\mathbf{x}$ is given by the Einstein-Smoluchowski theory:

$$P(\mathbf{x}, V|\mathbf{x}_0, \infty)d^n\mathbf{x} = C \exp\left(\frac{S(\mathbf{x}, \mathbf{x}_0)}{k_B}\right)d^n\mathbf{x} \tag{7.20}$$

where C is a normalization constant, k_B is Boltzmann's constant and $S(\mathbf{x}, \mathbf{x}_0)$ is the total entropy of the system having the extensive densities $\mathbf{x}_0$, containing the subsystem of finite volume V in the state $\mathbf{x}$. This expression applies to small fluctuations (i.e. for large values of the volume V). When the volume is infinite, the fluctuations vanish. Consequently:

$$P(\mathbf{x}, V|\mathbf{x}_0, \infty)d^n\mathbf{x} = \delta(\mathbf{x} - \mathbf{x}_0) \tag{7.21}$$

For large volumes, (7.21) leads to the Gaussian approximation:

$$\begin{aligned} P(\mathrm{x},\, V|\mathrm{x}_0,\, \infty) &= C \exp\left(\frac{V}{2k_B}\sum_{i,k}\frac{\partial^2 s(\mathrm{x}_0)}{\partial x_i\, \partial x_k}(x-x_0)_i\,(x-x_0)_k\right) \\ &= C \exp\left(\frac{1}{2k_B}\sum_{i,k} g^{ik}(\mathrm{X}_0)(x-x_0)_i\,(x-x_0)_k\right) \end{aligned} \tag{7.22}$$

One sees that the exponent in (7.22) is proportional to the square of the statistic (or thermodynamic) distance, measured from the equilibrium value. Therefore, this length element is the natural measure of the size of the fluctuations.

The important comment of Ruppeiner (1995) was that, although (7.20) depends on the choice of the parameters **x** used for defining the state, by mean of the volume $d^n\mathbf{x}$, the approximation (7.22) is invariant to the reparametrization. This makes the result have physical significance. One may conclude that this invariance must be kept for small volumes, too.

One starts from the thermodynamic limit $V \to \infty$ and the state $\mathbf{x}_0$. When the volume of the subsystem gets smaller, its fluctuations depend on the state of the neighborhoods. A Markov process is obtained for the fluctuations inside fluctuations. The associated Chapman-Kolmogorov equation describes how the fluctuation in a system of volume V depends on the state of the system at a slightly higher volume $V' = V + dV$. The role of the time is played by the quantity $1/V$, which becomes zero when the distribution function is given by delta function (7.21). With decreasing volume, the distribution takes smaller values, being given by (7.22). The Chapman-Kolmogorov equation takes the form of a Fokker-Planck covariant equation. In case the parameters are densities of extensive quantities, this last equation has the form:

$$\frac{\partial}{\partial V^{-1}} P(\mathbf{x},\, V|\mathbf{x}_0, \infty) = \frac{1}{k_B}\sum_{i,k}\left(\frac{\partial^2}{\partial x_i \partial x_k} g^{ik}(\mathbf{x}) P(\mathbf{x},\, V|\mathbf{x}_0, \infty)\right) \tag{7.23}$$

The equation complies with all conservation laws. These issues will be taken up and developed in subsequent chapters.

7.6 The Entropy Production

At macroscopic level, the Riemann structure is intimately related with the production of entropy. Let $\mathbf{X} = (U, V, N_1, N_2, \cdots) \in R^{n+1}$ and $\tilde{\mathbf{X}} = (\tilde{U}, \tilde{V}, \tilde{N}_1, \tilde{N}_2, \cdots) \in R^{n+1}$ be the vectors of the extensive variables of the systems

A and $\widetilde{A}$, respectively. During an interaction between the two systems, when the infinitesimal vector of flows $d\mathbf{X}$ passes from $\tilde{A}$ to A, the generation of entropy is:

$$dS_u = dS_A + dS_{\tilde{A}} = \sum_{i=1}^{n} (Y_i - \tilde{Y}_i) dX_i \tag{7.24}$$

where

$$\mathbf{Y} = \frac{\partial S}{\partial \mathbf{X}} \equiv \left(\frac{1}{T}, \frac{P}{T}, \frac{\mu_1}{T}, \frac{\mu_2}{T}, \cdots \right) \tag{7.25}$$

and the following conservation law was used

$$dX_i = -d\tilde{X}_i \qquad (i = 1, \cdots, n) \tag{7.26}$$

Equation (7.24) is the common expression, describing the entropy production as a product between a thermodynamic flow (or thermodynamic flux) and a thermodynamic force. It resembles the length element dl^2 of Riemann geometry. To show this, the expression (7.24) is re-written in the following form:

$$dS_u = -\sum_{i=1}^{n} \Delta Y_i dX_i = -\Delta \mathbf{Y} \cdot d\mathbf{X} \tag{7.27}$$

where $\Delta \mathbf{Y} = \tilde{\mathbf{Y}} - \mathbf{Y}$. The second law of thermodynamics asks that the sum in (7.27) must be positive. The element of thermodynamic length may be written as:

$$dl^2 = \sum_{j=1}^{n} \sum_{i=1}^{n} g(\mathbf{X}) dX_i dX_j \tag{7.28}$$

which is a symmetric product. Because:

$$dY_j = \sum_{i=1}^{n} \frac{\partial Y_j}{\partial X_i} dX_i = \sum_{i=1}^{n} \frac{\partial^2 S}{\partial X_i \partial X_j} dX_i = -\sum_{i=1}^{n} g^{ij}(\mathbf{X}) dX_i \tag{7.29–31}$$

the length element can be written now as:

$$dl^2 = -d\mathbf{Y} \cdot d\mathbf{X} \tag{7.32}$$

Note the similarity between (7.32) and (7.27). Also, although the terms $d\mathbf{Y}$ and $\Delta \mathbf{Y}$ look similar, they represent quantities of different nature. $d\mathbf{Y}$ is an infinitesimal change in the system state A, while $\Delta \mathbf{Y}$ is the difference $\tilde{\mathbf{Y}} - \mathbf{Y}$.

The equivalence of the metrics G and $g(\mathbf{X})$ is easily proved by using the symmetric product $-d\mathbf{Y}\cdot d\mathbf{X}$ of (7.32). Substituting:

$$dX_1 = dU = TdS + \sum_{i=2}^{n} W_i dZ_i \tag{7.33–35}$$

$$W_i = \frac{\partial U}{\partial Z_i} = -TY_i, \quad Z_i = X_i, \quad (i = 2, \cdots, n)$$

in (7.32) one obtains:

$$dl^2 = -d\mathbf{Y}\cdot d\mathbf{X} = \frac{d\mathbf{W}\cdot d\mathbf{Z}}{T} = \frac{G^{ij}dZ_i dZ_j}{T} = \frac{dL^2}{T} \tag{7.36–39}$$

Equations (7.36)–(7.39) represent the Gouy-Stodola theorem in infinitesimal form. This theorem shows the connection between the loss availability (or exergy) at temperature T and the associated entropy generation.

7.7 Dissipation for Near Equilibrium Processes

There is a particularly important case for which there is a close connection between $d\mathbf{Y}$ and $\Delta\mathbf{Y}$. This case corresponds to the equilibration of a system in contact with a thermal bath (i.e. in contact with a thermodynamic system whose temperature can be controlled).

Consider a system large enough that any change in its intensive variables $\tilde{\mathbf{Y}}$ can be neglected. It is allowed to the system A to equilibrate with the system $\tilde{A}$, so that the final values of $\mathbf{Y}$ to be equal with $\tilde{\mathbf{Y}}$. The integration of the entropy generation on the path toward equilibrium gives:

$$\Delta S_u = \int dS_u = -\int \Delta\mathbf{Y}\cdot d\mathbf{X} \tag{7.40}$$

which, in virtue of (7.31) becomes:

$$\Delta S_u = \int \Delta\mathbf{X}^t g(\mathbf{X}) d\mathbf{X} \tag{7.41}$$

up to the first order of approximation in $\Delta\mathbf{X} = \mathbf{X}_0 - \mathbf{X}$, where $\mathbf{X}_0$ is the vector of the extensive variables of the system A, after equilibrating the bath $\tilde{A}$. In this approximation order, one may assume that the matrix of the metric $g(\mathbf{X})$ is constant. Then (7.41) can be integrated, resulting in:

$$\Delta S_u = \frac{1}{2}\Delta \mathbf{X}^t g(\mathbf{X})\Delta \mathbf{X} = \frac{1}{2}\Delta l^2 \tag{7.42, 43}$$

7.8 Discrete Theorems of Horse-Carrot Kind

The Horse-carrot theorems are resulting from the general expression (7.43), which gives the connection between the length of an equilibration process and the corresponding entropy generation.

Consider a path in the space of the states of system A and an equilibration process. Select the states of k thermal baths, used to match the values of the intensive variable of the system $\bar{Y}$ in k points along the path.

The discrete horse-carrot theorem answers the question: how the states of the thermal baths must be chosen so as to minimize the total generation of entropy in the k successive equilibrations of the k baths, along the path?

For large k numbers, the answer is simple: the baths should be placed equidistantly in the geometry described by $g(\mathbf{X})$. This conclusion comes from the observation that the overall entropy generation in the k processes, which is given by

$$\Delta S_u = \sum_{j=1}^{k} \Delta_j S_u = \frac{1}{2}\sum_{j=1}^{k} \Delta_j l^2 \tag{7.44}$$

must be minimized, with the restriction that the total length of the k processes is fixed:

$$l = \sum_{j=1}^{k} \Delta_j l \tag{7.45}$$

Minimization of entropy generation is done by using the Lagrange multiplier method and this yields:

$$\Delta_j l = \text{constant} = \frac{l}{k} \tag{7.46}$$

Substituting this result in (7.44) one obtains the so-called horse-carrot inequality:

$$\Delta S_u \geq \Delta S_u^{\min} = \frac{l^2}{2k} \tag{7.47}$$

A more detailed analysis shows that, for large times, the same number of relaxation times should be allocated for each equilibration process. So, the path of minimum entropy generation, to drive the system in a finite time, along a given path, by using a fixed number k of intermediate equilibrations, is of making equidistant steps, with a constant number of relaxation times allocated for each step.

This is the origin of introducing the concept of *thermodynamic speed,* defined by:

$$v \equiv \frac{dl}{d\xi} \tag{7.48}$$

where $d\xi = dt/\varepsilon$, t is the time and ε is the relaxation time of the system.

Therefore, for a process in k steps, the constant thermodynamic speed is optimal.

7.9 Comments on the Loss of Exergy

An expression fully similar to Eq. (7.43) can be obtained for the loss of exergy ΔA_u during a small equilibration process:

$$\Delta A_u = \frac{1}{2}\Delta \mathbf{Z}^t G \Delta \mathbf{Z} = \frac{1}{2}\Delta L^2 \tag{7.49, 50}$$

By analyzing these relationships one can understand why the conformal equivalence between the two metrics given by the seconds derivatives of S and U is of primary importance. This conformal equivalence, expressed by Eq. (7.38), is the differential form of the Gouy-Stodola theorem:

$$\Delta A_u = -T_a \Delta S_u \tag{7.51}$$

where T_a is the temperature of the "environment". For the infinitesimal process (7.38), the role of the "environment" is played by the system $\tilde{A}$.

The ambiguity about the place where the heat associated to the exergy loss ΔA_u is delivered, severely limits the usefulness of the following horse-carrot type inequality:

$$\Delta A_u \geq \frac{L^2}{2k} \tag{7.52}$$

For isothermal processes, the inequalities (7.47) and (7.52) are equivalent. For non-isothermal processes, however, the thermodynamic lengths l and L, respectively, are not linked by a simple relationship. Although the length L is sometimes useful, dL^2 is often used for the assessment of dl^2, through the relationship (7.38). This will be discussed in the next section.

7.10 A Continuous Theorem of Horse-Carrot Kind

Consider the following problem. The thermodynamic system A must follow the path $\mathbf{X}(t)$, $t \in [0, \tau]$. Which is the amount of entropy generated during the evolution?

If the total duration τ of the process is high enough, the answer is similar to that found in case of discrete processes. This time the optimal control analysis leads to the conclusion that the optimal control means a constant rate of entropy generation, instead of the prior conclusion, which was that of the optimal control means a constant thermodynamic speed.

It is assumed that one can control in a reversible way the system $\tilde{A}$ and that the system A is only indirectly affected, through its contact with the system $\tilde{A}$. The reasoning starts from the integral form of the relationship (7.27) for the total generation of entropy:

$$\Delta S_u = \int_0^\tau dS_u = -\int_0^\tau \Delta \mathbf{Y} \cdot d\mathbf{X} \tag{7.53}$$

Equation (7.53) can be rearranged to highlight the link between dissipation and geometry. Define a state $\mathbf{X}_e$ by the following simple relationship:

$$-\left(\tilde{\mathbf{Y}} - \mathbf{Y}\right) = g(\mathbf{X})(\mathbf{X}_e - \mathbf{X}) \tag{7.54}$$

The metric $g(\mathbf{X})$ is not necessarily invertible. In case that $\mathbf{X}_e$ consists of several homogenous phases, it is necessary that their scales are separated. This condition can be accomplished by specifying the way each of these scales of A evolves.

Consider a state $\mathbf{X}_e$ close to state $\mathbf{X}$. In this case, the state $\mathbf{X}_e$ can be interpreted as a state of the system A, which minimizes the entropy production rate, when this system is in contact with the present state of the system $\tilde{A}$, provided that A is on the path though $\mathbf{X}$ on the direction $d\mathbf{X}$.

Substituting (7.54) in (7.53) and replacing $d\mathbf{X}$ by $(d\mathbf{X}/dl)dl$, one finds the total production of entropy:

$$\Delta S_u = \int_0^\tau (\mathbf{X}_e - \mathbf{X})^t g(\mathbf{X}) \frac{d\mathbf{X}}{dl} dl \tag{7.55}$$

Remind that a metric defines an inner product on a vector space. The integrand in Eq. (7.55) is the inner product obtained using the metric $g(\mathbf{X})$ of the deviation $\mathbf{X}_e - \mathbf{X}$ and the unit tangent vector $d\mathbf{X}/dl$. Therefore, this integrand can be interpreted as the distance between the current state and the state that the system is trying to achieve, projected on the direction of $d\mathbf{X}$. Denote by $D = dS_u/dl$ this projected distance. Note that by using a mean theorem, the total dissipation can be written as:

$$\Delta S_u = \int_0^l D dl = \bar{D} l \tag{7.56}$$

So, the total generation of entropy is given by the product of the average distance to equilibrium and the total distance traveled.

Another interesting relationship can be found by considering the quantity:

$$\varepsilon = \frac{dS_u/dt}{(dl/dt)^2} \tag{7.57}$$

Keep in mind that ε has dimension of time. It is also seen that for sufficiently slow processes with separable time scales, ε is just the relaxation time. This is seen by writing:

$$\frac{d\mathbf{X}}{dt} = \frac{\mathbf{X}_e - \mathbf{X}}{\varepsilon} \tag{7.58}$$

According to the definition of $\mathbf{X}_e$, $d\mathbf{X}/dt$ and $\mathbf{X}_e - \mathbf{X}$ should have the same direction and therefore they are proportional. It is assumed that the process is slow and the time scales are separable. Then, except the slowest mode of the system, the other modes must balance instantly. Therefore, $\mathbf{X}_0 - \mathbf{X}$ must be proportional with $d\mathbf{X}/dt$, i.e. $\mathbf{X}_e = \mathbf{X}_0$.

With or without accepting the hypotheses of slow process and separable time scales, respectively, the definition (7.57) allows writing the total generation of entropy as follows:

$$\Delta S_u = \int_0^\tau \varepsilon \frac{d\mathbf{X}^t}{dt} g(\mathbf{X}) \frac{d\mathbf{X}}{dt} dt = \int_0^\tau \varepsilon \left(\frac{dl}{dt}\right)^2 dt \tag{7.59}$$

Applying a mean theorem, one finds:

$$\Delta S_u = \bar{\varepsilon} \int_0^\tau \frac{d\mathbf{X}^t}{dt} g(\mathbf{X}) \frac{d\mathbf{X}}{dt} = \bar{\varepsilon} \int_0^\tau \left(\frac{dl}{dt}\right)^2 dt \tag{7.60}$$

A third interesting expression of ΔS_u is obtained if the parameter along the trajectory $\mathbf{X}(t)$ is changed and the dissipation integral is expressed in terms of the number of relaxations, ξ. Note that $d\xi = dt/\varepsilon$. Thus:

$$\Delta S_u = \int_0^{\Xi} \frac{d\mathbf{X}^t}{d\xi} g(\mathbf{X}) \frac{d\mathbf{X}}{d\xi} d\xi = \int_0^{\Xi} \left(\frac{dl}{d\xi}\right)^2 d\xi \tag{7.61}$$

where

$$\Xi = \int_0^{\tau} d\xi = \int_0^{\tau} \frac{dt}{\varepsilon} \tag{7.62}$$

All three previously deducted relationships are generally valid. However, the physical significance of D and ε is formal, except for the case when the processes are slow and the time scales are separable, when A and $\tilde{A}$ are almost in equilibrium at any time.

A fourth relationship involving ΔS_u is obtained from a linearized flux-thermodynamics force equation of Onsager–Prigogine type:

$$\frac{d\mathbf{X}}{dt} = \gamma \cdot \Delta\mathbf{Y} \tag{7.63}$$

where γ is the kinetic coefficients matrix (which is symmetric and positively defined). If the Eq. (7.63) is solved for $\Delta\mathbf{Y}$ and replaced in (7.53), one obtains:

$$\Delta S_u = \int_0^{\tau} \frac{d\mathbf{X}^t}{d\xi} \gamma^{-1} \frac{d\mathbf{X}}{dt} dt = \int_0^{\tau} \left(\frac{d\lambda}{dt}\right)^2 dt \tag{7.64}$$

In this expression one sees the integral of a squared speed. This time λ is given by another metric (i.e. by γ^{-1}). There is a fundamental difference between this metric and the other metrics previously considered. The coefficients in the metric γ are kinetic quantities, unlike the coefficients entering the previously considered metrics, which are equilibrium quantities. In other words, the coefficients of the metric $g(\mathbf{X})$ are covariant quantities while the coefficients of γ are temporal correlations.

7.11 A Simple Optimization Lemma

In some of the cases discussed in the previous sections, an integral of the following form it minimized:

$$\int_0^\tau f(x)\left(\frac{dx}{dt}\right)^2 dt \tag{7.65}$$

with given values for $x(0)$ and $x(\tau)$. Imposing the necessary Euler-Lagrange conditions for the autonomous Lagrangian K:

$$K = f(x)\left(\frac{dx}{dt}\right)^2 \tag{7.66}$$

one obtains (see Badescu 2003):

$$K - \frac{dx}{dt}\frac{\partial K}{\partial(dx/dt)} = \text{const.} \tag{7.67}$$

Substituting the expression of K from (7.66) into (7.67), one finds that, to achieve optimality, the Lagrangian K must be constant.

A useful corollary corresponds to the particular case $f = 1$:

$$K = \left(\frac{dx}{dt}\right)^2 = \text{const.} \tag{7.68}$$

This implies

$$\frac{dx}{dt} = \text{const.} = \frac{\Delta x}{\tau} \tag{7.69}$$

and the extreme value of the integral will be:

$$\int_0^\tau \left(\frac{dx}{dt}\right)^2 dt \geq \frac{\Delta x^2}{\tau} \tag{7.70}$$

This particular case is of special importance when x is the length of the arc associated with a certain metric M. The coordinates in this space are denoted $\mathbf{X}$. Then, the inequality (7.70) becomes:

$$\int_0^\tau \frac{d\mathbf{X}^t}{dt} M \frac{d\mathbf{X}}{dt} dt \geq \left(\int_0^\tau \left(\frac{d\mathbf{X}^t}{dt} M \frac{d\mathbf{X}}{dt}\right)^{1/2} dt\right)^2 / \tau \tag{7.71}$$

7.11.1 Lemma Applications

The application of the previous Lemma for relations (7.59) and (7.64) lead to the conclusion that, in order to minimize the entropy production, one must use a constant rate of entropy generation:

$$\dot{S}_u = \varepsilon \left(\frac{dl}{dt} \right)^2 = \left(\frac{d\lambda}{dt} \right)^2 = \text{const.} \tag{7.72}$$

Applying the corollary of that Lemma to the squared speed in relationships (7.60) and (7.61) one obtains different continuous versions of the horse-carrot inequality:

$$\Delta S_u \geq \frac{\bar{\varepsilon} l^2}{\tau}, \quad \Delta S_u \geq \frac{l^2}{\Xi} \tag{7.73, 74}$$

These inequalities impose lower limits for the dissipations occurring during the processes, namely the square of the thermodynamic distance over the number of relaxations. The inequalities (7.73) and (7.74) resemble the inequality (7.47). However, these inequalities do not say anything useful about how to monitor the entropy generation in a given time interval. Indeed, the averaging process in which $\bar{\varepsilon}$ and Ξ are obtained for a given interval of time, is based on an implicit dependence.

References

Badescu, V.: Control optimal in ingineria proceselor termice. Matrix Rom, Bucuresti (2003)

Diosi, L., Salamon, P.: From statistical distances to minimally dissipative processes. In: Sieniutycz, S., De Vos, A. (eds.) Thermodynamics of Energy Conversion and Transport, pp. 286–318. Springer, New York (2000)

Ruppeiner, G.: Thermodynamics: a Riemanian geometric model. Phys. Rev. A **20**, 1608 (1979)

Ruppeiner, G.: Riemannian geometry in thermodynamic fluctuation theory. Rev. Mod. Phys. **67**, 605–659 (1995)

Weinhold, F.: Metric geometry of equilibrium thermodynamics. J. Chem. Phys. **63**, 2479 (1975)

Chapter 8
Geometrization of Thermodynamic Fluctuations

The entropy S can be calculated based on the number of microstates Ω of a thermodynamic system by using the well-known formula due to Ludwig Boltzmann:

$$S = k_B \ln \Omega \tag{8.1}$$

The inverse of this formula is:

$$\Omega = \exp\left(\frac{S}{k_B}\right) \tag{8.2}$$

Einstein made from the relationship (8.2) a starting point of the theory of thermodynamic fluctuations. In this chapter, the excellent presentation by Ruppeiner (1995) is followed closely. Further details may be found in Ruppeiner (1979, 1991).

Consider a closed thermodynamic system A_{V_0}, very large, in equilibrium. This system has fixed volume V_0 and fixed energy per unit volume u_0. It is also considered a subsystem A_V of A_{V_0}. The energy per unit volume of A_V is denoted by u and is not subjected to restrictions. All other thermodynamic parameters of the system A_V are constant.

8.1 Fluctuations of a Thermodynamic Quantity (Classical Theory)

In the usual formulation, the second law of thermodynamics states that, at equilibrium, the internal energy u has a specific value that maximizes the entropy $S_0(u, u_0)$ of the system A_{V_0}. Statistical mechanics brings more information, allowing description of the fluctuations around this state of maximum entropy. The statistical method will be exemplified further by using the microcanonic ensemble. The basic

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_8

postulate is that all accessible microstates of A_{V_0} have the same probability of occurring. Therefore, the probability of finding the internal energy of the subsystem A_V between u and $u + du$ is proportional to the number of microstates of A_{V_0} corresponding to this interval:

$$P_V(u|u_0)du = C\Omega_V(u, u_0)du \tag{8.3}$$

where $\Omega_V(u, u_0)$ is the density of states and C is a normalization constant. Using (8.2) and (8.3) yields:

$$P_V(u|u_0)du = C\exp\left[\frac{S_0(u, u_0)}{k_B}\right]du \tag{8.4}$$

8.1.1 Gaussian Approximation

The next step is to consider the limit $V_0 \to \infty$ and to approximate the probability density $P(u|u_0)$ by developing the entropy around its maximum at $u = u_0$ and keeping only terms lower than the second order. After elementary calculations, one finds:

$$P_V(u|u_0)du = \left(\frac{Vg(u_0)}{2\pi}\right)^{1/2}\exp\left[-\frac{V}{2}g(u_0)(u - u_0)^2\right]du \tag{8.5}$$

where

$$g(u_0) \equiv -\frac{1}{k_B}\left(\frac{d^2s}{du^2}\right)_{u=u_0} \tag{8.6}$$

and $s = s(u)$ is the entropy per unit volume in the thermodynamic limit. Relation (8.5) is the so-called *Gaussian approximation*, which is valid for small deviations of u with respect to u_0.

8.1.2 Difficulties of the Gaussian Approximation

The Gaussian approximation, despite its simplicity, hides some conceptual difficulties. To illustrate this, assume that instead of the energy density u, another thermodynamic parameter $x = x(u)$ is used for characterizing the thermodynamics state of the system A_V.

The counting of the microstates that led to Eq. (8.3) can be repeated in case of the parameter x. Assume, again, that the entropy is proportional to the logarithm of the states density. Then, the result should be the same as that given by Eq. (8.4), with x replacing u. However, if one starts with the Eq. (8.4) and the change of coordinates $u \rightarrow u(x)$ is performed, a different result is obtained. The left-hand side of (8.4) becomes:

$$P_V(u|u_0)du = \left[P_V(u|u_0)\left[\frac{du}{dx}\right]\right]dx \equiv P_V(x|x_0)dx \tag{8.7}$$

which gives the probability of finding the new parameter in the interval x to $x + dx$. In the right-hand side of (8.4), the entropy is a state function, invariant with respect to the coordinate transformation. Thus:

$$P_V(x|x_0)dx = C\left[\frac{du}{dx}\right] \exp\left[\frac{S_0(x, x_0)}{k_B}\right]dx, \tag{8.8}$$

which does not have the form of Eq. (8.4), since the derivative of u from the right-hand side is not generally constant and, therefore, cannot be absorbed in the normalization constant C.

The conclusion is that, when using statistical mechanics one obtains an expression which is different from that obtained by using purely mathematical reasoning. The last type of reasoning clearly shows that the thermodynamic fluctuations formula (8.4) depends on coordinates. It is said that this equation is not covariant.

Also, the Eq. (8.4) suffers of a certain inconsistency. To show this, define the average value of a thermodynamic function as:

$$<f> \ = \int f(x)P_V(x|x_0)dx \tag{8.9}$$

On the other hand, there is no principle allowing to state that the entropy function, which is involved indirectly in the average definition (8.9), by means of the relation (8.8), guarantees the relationship:

$$<u> \ = u_0 \tag{8.10}$$

when $V_0 \rightarrow \infty$. From the physical point of view, Eq. (8.10) must be valid for any value of V, but it is clear that when using (8.4) this does not happen, unless $S_0(u, u_0)$ is an even function of $u - u_0$. No law of thermodynamics guarantees this.

8.1.3 Advantages of the Gaussian Approximation

Even with the drawbacks mentioned above, the Gaussian approximation (8.5) has a number of features that make it appealing.

First, it is becoming increasingly more accurate when the volume V of the subsystem gets bigger, because in this case the fluctuations of the intensive quantities are becoming smaller.

Second, the Gaussian approximation is covariant. A coordinate transformation $u \rightarrow x(u)$ in Eq. (8.5) shows that the second-order approximation of the displacements with respect to $x_0(=x(u_0))$ is:

$$P_V(x|x_0)du = \left(\frac{Vg(x_0)}{2\pi}\right)^{1/2} \exp\left[-\frac{V}{2}g(x_0)(x-x_0)^2\right]dx \tag{8.11}$$

where

$$g(x_0) \equiv \left[\frac{du}{dx}\right]^2 g(u_0) \tag{8.12}$$

Since the fluctuations are small in the range of validity of the Gaussian approximation, it is expected that the derivative in (8.12) is nearly constant in the region of the states with reasonable probability. Therefore, the Eq. (8.11) has exactly the same form as the Eq. (8.5), which is written in energy coordinates.

Third, from the symmetry of the Gaussian distribution, it is seen that the average value of x is x_0, regardless of the choice of the coordinate, which could be the energy density.

8.2 Classical Fluctuation of Two Independent Quantities

The above discussion refers to the case when just one single thermodynamic quantity fluctuates. The fluctuations of two independent variables will be treated in the following. As an example, consider a subsystem that exchange particles and energy with the environment. If there are two variables, it is possible to develop a formalism based on Riemann geometry, leading to a non-zero thermodynamic curvature.

The basic concept of a two-dimensional Riemann geometry is a surface, or a two-dimensional manifold that is, roughly speaking, a set of points smoothly parametrized by two coordinates, $x = (x^0, x^1)$. As shown in the Chap. 1, the points on the manifold represent physical quantities with intrinsic meaning (here, they are thermodynamic states).

The second element of the Riemann geometry is a procedure that defines the distance Δl for every pair of neighboring points, having the coordinate differences Δx^α. The distance is given by the quadratic form:

$$(\Delta l)^2 = \sum_{\mu,\nu=0}^{1} g_{\mu\nu}(x)\Delta x^\mu \Delta x^\nu \tag{8.13}$$

where the matrix of the coefficients $g_{\alpha\beta}(x)$ is called the metric tensor. One reminds that a manifold where the distance has been defined by using the relationship (8.13) is called Riemann manifold. A key requirement is that, for a given couple of points, the distance Δl must be independent of the coordinate system used to specify those points.

It is important to note that there is no requirement that these two-dimensional Riemannian manifolds are surfaces which can be embedded in three-dimensional flat spaces, although the latter are of course examples of Riemannian manifolds.

To represent a physical problem using a Riemann manifold, it is necessary, first, to have a set of physical objects corresponding to the set of the points on the manifold, and, second, that the distance, which is defined by a quadratic form of type (8.13), has a physical motivation and is unique.

Approaching the theory of thermodynamic fluctuations in the case of two parameters requires a generalization of the case discussed in Sect. 8.1, by considering an open subsystem A_V of the closed infinite system A_{V_0}. It is allowed the exchange of particles between subsystem and environment, so that the energy per unit volume a^0 and the number of particles per unit volume can fluctuate.

Developing the total entropy in Taylor series around its maximum, up to second order terms, leads to the following expression (Ruppeiner 1995):

$$P_V(a|a_0)da^0da^1 = \frac{V}{2\pi}\exp\left[-\frac{V}{2}\sum_{\mu,\nu=0}^{1} g_{\mu\nu}(a_0)\Delta a^\mu \Delta a^\nu\right](g(a_0))^{\frac{1}{2}}da^0da^1 \qquad (8.14)$$

where $\Delta a^\alpha = a^\alpha - a_0^\alpha$ and

$$g_{\alpha\beta}(a_0) = -\frac{1}{k_B}\frac{\partial^2 s}{\partial a^\alpha \partial a^\beta}, \quad g(a_0) = \det[g_{\alpha\beta}(a_0)] \qquad (8.15, 16)$$

For two given neighboring states, the value of the positive defined quadratic form, given by

$$(\Delta l)^2 = \sum_{\mu,\nu=0}^{1} g_{\mu\nu}(a_0)\Delta a^\mu \Delta a^\nu \qquad (8.17)$$

is independent of the choice of coordinate system, since the difference of entropy between two states does not depend on the coordinate system used to represent those states. This is a necessary condition, because the probability of a fluctuation depends only on the thermodynamic states.

Note that the unit of the square of the thermodynamic length $(\Delta l)^2$ is that of an inverse volume.

8.3 Classical Gaussian Approximation in the General Case

One can discern two categories of parameters that can be used to specify the thermodynamic state of a subsystem. First, there are mechanical parameters such as energy, number of particles, magnetization, which can be calculated by adding the corresponding microscopic quantities. The second category consists of thermodynamic parameters, such as temperature, chemical potential and entropy. They are in connection with the distributions over microscopic states, and, unlike the mechanical parameters, they have no equivalent in microscopic quantities. The former parameters are most suitable for describing the thermodynamic state of a subsystem.

Consider an open subsystem A_V, of fixed volume V, of a closed thermodynamic system A_{V_0} of a very large volume V_0. The system A_{V_0} consists of r fluid components, supposed not to interact chemically but being in equilibrium. Denote by $a_0(a_0^1, a_0^2, \ldots, a_0^n)$ the group of quantities consisting of the internal energy per unit volume and the number of particles per unit volume, corresponding to the r components of A_{V_0}. These mechanical parameters represent the standard densities in the entropy representation and constitute the thermodynamics state of A_{V_0}. Similarly, the subsystem A_V has the thermodynamic state a. The classical theory of fluctuations is based on the following three axioms.

Axiom 8.1 A_V and its reservoir $A_{V_c} = A_{V_0} - A_V$, having the volume $V_c = V_0 - V$, are homogeneous systems, which, in the thermodynamic limit, can be described thermodynamically.

Axiom 8.2 The conditional probability of finding A_V in a state for which the parameters are found between a and $a + da$, provided that A_{V_0} is found in the state a_0, is

$$P_V(a|a_0)da^0da^1\cdots da^r = C\exp\left[\frac{S(a,a_0)}{k_B}\right]da^0da^1\cdots da^r \tag{8.18}$$

where $S(a, a_0)$ is the entropy of the system A_{V_0} when the subsystem A_V is found in the state a, and C is a normalization constant.

Axiom 8.3 The entropy is additive but is not conserved. The standard extensive parameters (internal energy and number of particles) are additives and conserved.

The next step consists of developing the entropy in series up to second order terms around the homogenous state, where $a = a_0$. Using the Axioms 8.1 and 8.3, one finds:

$$S(a, a_0) = Vs(a) + V_cs(a_c) \tag{8.19}$$

where $s(a)$ is the entropy per unit of volume and a_c are the standard densities of A_{V_c}. Using Axiom 8.3, one can write:

$$a_c^\alpha = \frac{V_0 a_0^\alpha - V a^\alpha}{V_c} \tag{8.20}$$

Developing the entropy densities in (8.19) around $a_c^\alpha = a_0^\alpha = a^\alpha$, it is obtained:

$$\begin{aligned} S(a, a_0) = {} & V s(a_0) + V_c s(a_0) + V \frac{\partial s}{\partial a^\mu}(a^\mu - a_0^\mu) + V_c \frac{\partial s}{\partial a^\mu}(a_c^\mu - a_0^\mu) \\ & + \frac{1}{2} V \frac{\partial^2 s}{\partial a^\mu \partial a^\nu}(a^\mu - a_0^\mu)(a^\nu - a_0^\nu) \\ & + \frac{1}{2} V_c \frac{\partial^2 s}{\partial a^\mu \partial a^\nu}(a_c^\mu - a_0^\mu)(a_c^\nu - a_0^\nu) + \cdots \end{aligned} \tag{8.21}$$

where all derivatives are evaluated in a_0.

The Einstein summation convention is used in the following. It means that for repeated indices in products, all contributions are summed. Also, it is accepted the convention that the list of the indices which do not contribute to the summation in an expression begins with α while the list of the indices which do contribute to the summation begins with μ.

Substituting (8.20) in (8.21) and passing to the limit $V_c \to \infty$, one obtains:

$$S(a, a_0) = V_0 s(a_0) + \frac{1}{2} V \frac{\partial^2 s}{\partial a^\mu \partial a^\nu} \Delta a^\mu \Delta a^\nu + \cdots \tag{8.22}$$

where

$$\Delta a^\alpha \equiv a^\alpha - a_0^\alpha \tag{8.23}$$

Cancellation of the first order terms in the development shows that the state $a = a_0$ corresponds to an extreme of the total entropy. By imposing the usual conditions of stability, it can show that the extreme is a maximum. Alternately, the stability condition may be imposed by asking that the entropy has a maximum for $a = a_0$.

Equation (8.22) is valid only in case of standard densities (or in the case of linear transformations of these densities), because these coordinates have the additivity property. The general case, in which S_0 is described in an arbitrary system of thermodynamic coordinates $x = x(a)$, requires a careful analysis. It may be shown that the entropy development up to terms of second order around the maximum of $x = x_0 \equiv x(a_0)$ leads to (Ruppeiner 1995):

$$S(x, x_0) = S_0(x_0, x_0) + \frac{1}{2} \frac{\partial^2 S_0}{\partial x^\mu \partial x^\nu} \Delta x^\mu \Delta x^\nu + \cdots \tag{8.24}$$

where $\Delta x^{\alpha} \equiv x^{\alpha} - x_0^{\alpha}$. Here, unlike the situation in (8.22), the second derivatives are not typical thermodynamic derivatives, because $S_0(x, x_0)$ depends on two distinct thermodynamic systems, with different states. These derivatives are now expressed in terms of thermodynamic quantities. First, they have to be transformed into another coordinate system $x' = x'(x)$. Using the common rule:

$$\frac{\partial}{\partial x^{\alpha}} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial}{\partial x'^{\mu}} \tag{8.25}$$

it may be shown that

$$\frac{\partial^2 S_0}{\partial x^{\alpha} \partial x^{\beta}} = \frac{\partial^2 x'^{\mu}}{\partial x^{\alpha} \partial x^{\beta}} \frac{\partial S_0}{\partial x'^{\mu}} + \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \frac{\partial^2 S_0}{\partial x'^{\mu} \partial x'^{\nu}} \tag{8.26}$$

At the maximum of S_0, this expression becomes

$$\frac{\partial^2 S_0}{\partial x^{\alpha} \partial x^{\beta}} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \frac{\partial^2 S_0}{\partial x'^{\mu} \partial x'^{\nu}} \tag{8.27}$$

which represents the transformation rule for the components of a tensor of the second rank. Now it is possible to express the general second order derivatives in terms of thermodynamic quantities. By comparing Eqs. (8.22) and (8.27) it is found that

$$\frac{\partial^2 S_0}{\partial x^{\alpha} \partial x^{\beta}} = V \frac{\partial a^{\mu}}{\partial x^{\alpha}} \frac{\partial a^{\nu}}{\partial x^{\beta}} \frac{\partial^2 s}{\partial a^{\mu} \partial a^{\nu}} \tag{8.28}$$

Using the quadratic development, the Gaussian approximation of the theory of thermodynamic fluctuations can be written as follows:

$$\begin{aligned} & P_V(x|x_0) dx^0 dx^1 \cdots dx^r \\ & = \left[\frac{V}{2\pi}\right]^{n/2} \exp\left[-\frac{V}{2} g_{\mu\nu}(x_0) \Delta x^{\mu} \Delta x^{\nu}\right] (g(x_0))^{1/2} dx^0 dx^1 \cdots dx^r \end{aligned} \tag{8.29}$$

where

$$g_{\alpha\beta}(x_0) = -\frac{1}{V k_B} \left(\frac{\partial^2 S_0}{\partial x^{\alpha} \partial x^{\beta}}\right)_{x = x_0}, \quad g(x_0) = \det[g_{\alpha\beta}(x_0)] \tag{8.30, 31}$$

Using the probability distribution (8.29), the following average values can be calculated:

$$< \Delta x^{\alpha} > \, = 0, \quad < \Delta x^{\alpha} \Delta x^{\beta} > \, = \frac{g^{\alpha\beta}(x_0)}{V} \tag{8.32, 33}$$

where $g^{\alpha\beta}$ represent the elements of the inverse matrix of $g_{\alpha\beta}$.

One can easily verify that the quadratic form in the argument of the exponential (8.29) transforms itself as a scalar under a change of coordinates. In the first order of approximation

$$\Delta x^{\alpha} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \Delta x'^{\mu} \tag{8.34}$$

This is the transformation rule for a contravariant tensor of first rank. Substituting (8.34) in the quadratic form and using (8.27), with the terms noted prime and without prime interchanged, the expected invariance is revealed. This is necessary, because neither the probability of a fluctuation between two states, nor the entropy difference should not depend on the choice of the coordinate system used to describe those states.

The quadratic form

$$(\Delta l)^2 \equiv g_{\mu\nu}(x_0) \Delta x^{\mu} \Delta x^{\nu} \tag{8.35}$$

constitutes a positive definite Riemann metric, on the space of the thermodynamic states. It is independent of the volume V (see Eqs. (8.28) and (8.30)). The physical interpretation of the distance between two thermodynamic states is that the less is the probability of a fluctuation between the two states, the larger is the distance between the two states. Note that

$$(g(x_0))^{1/2} dx^0 dx^1 \cdots dx^r \tag{8.36}$$

is the invariant element of volume on a Riemann manifold.

The condition that the element of thermodynamics length Δl be positive for any set of values of the differentials of the coordinates impose conditions on the elements of the metric $g_{\alpha\beta}$. First, all elements on the diagonal of the metric must be positive. These elements represent, for example, heat capacities and compressibility factors. The condition that they are positive is the thermodynamic stability condition.

However, the fact that the diagonal elements are positive does not ensure the positivity of the element of thermodynamics length. A necessary and sufficient condition is that leading minors of the metric elements, defined by:

$$\begin{gathered} p_0 \equiv 1, \quad p_1 \equiv g_{00}, \quad p_2 \equiv \det\begin{pmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{pmatrix}, \\ p_3 \equiv \det\begin{pmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{pmatrix}, \ldots, p_n \equiv \det(g) \end{gathered} \tag{8.37a–e}$$

be all positive.

8.3.1 Particular Forms of Thermodynamic Metrics

Next it will be shown the form taken by the thermodynamics metric in different coordinate systems. Two useful relationships will be deduced on this basis. First, the following notation is adopted

$$F = \left(\frac{1}{T}, -\frac{\mu^1}{T}, \ldots, -\frac{\mu^r}{T}\right) \tag{8.38}$$

for the standard intensive quantities in the entropy representation:

$$F^\alpha = \frac{\partial s(a)}{\partial a^\alpha} \tag{8.39}$$

Here T is the temperature and μ^i is the chemical potential of the component i. The n quantities F^α constitute a complete set of coordinates in the phase space. The pressure p, which is the conjugate of the fixed volume V, is not expressed in terms of the other coordinates. Using (8.22) one finds:

$$(\Delta l)^2 = -\frac{1}{k_B}\Delta F^\mu \Delta^\mu \tag{8.40}$$

This relationship will be used to express the metric in different coordinate systems. Another useful relationship can be achieved on the basis of the variables used in the energy representation. In this case, s replaces u as the zeroth coordinate and the conjugate intensive variables are:

$$P = (T, \mu^1, \ldots, \mu^r) \tag{8.41}$$

Substituting

$$\begin{aligned} &\Delta a^0 = \Delta u = T\Delta s + \sum_{i=1}^{r} \mu^i \Delta a^i, \quad \Delta F^0 = -\frac{1}{T^2}\Delta T, \\ &\Delta F^i = \frac{\mu^i}{T^2}\Delta T - \frac{1}{T}\Delta\mu^i \end{aligned} \tag{8.42–44}$$

(with $1 \le i \le r$) in the relationship (8.40), one obtains

$$(\Delta l)^2 = \frac{1}{k_B T}\Delta T \Delta s + \frac{1}{k_B T}\sum_{i=1}^{r} \mu^i \Delta a^i \tag{8.45}$$

To express the metric in the F coordinates one writes

$$\Delta a^{\alpha} = \frac{\partial a^{\alpha}}{\partial F^{\mu}} \Delta F^{\mu} \tag{8.46}$$

and this expression is replaced in (8.40), yielding

$$(\Delta l)^2 = \frac{1}{k_B} \frac{\partial^2 \phi}{\partial F^{\mu} \partial F^{v}} \Delta F^{\mu} \Delta F^{v} \tag{8.47}$$

where

$$\phi(F^0, F^1, \ldots, F^r) = s - F^{\mu} a^{\mu} \tag{8.48}$$

It may be easily shown that

$$\phi = \frac{p}{T} \tag{8.49}$$

This thermodynamic potential is k_B/T times larger than the logarithm of the macrocanonic partition function.

To express the metric in the coordinates $(T, a^1, \ldots, a^r)$, one writes:

$$\Delta s = \left[\frac{\partial s}{\partial T}\right] \Delta T + \sum_{i=1}^{r} \left[\frac{\partial s}{\partial a^i}\right] \Delta a^i, \quad \Delta \mu^i = \left[\frac{\partial \mu^i}{\partial T}\right] \Delta T + \sum_{j=1}^{r} \left[\frac{\partial \mu^i}{\partial a^j}\right] \Delta a^i \tag{8.50a, b}$$

Using the Maxwell relationship:

$$\left[\frac{\partial s}{\partial \rho^i}\right] = -\left[\frac{\partial \mu^i}{\partial T}\right] \tag{8.51}$$

and substituting it in (8.45), one obtains

$$(\Delta l)^2 = \frac{1}{k_B T} \left[\frac{\partial s}{\partial T}\right] (\Delta T)^2 + \frac{1}{k_B T} \sum_{i,j=1}^{r} \left[\frac{\partial \mu^i}{\partial a^j}\right] \Delta a^i \Delta a^j \tag{8.52}$$

This element of thermodynamic length is diagonalized for the case $r = 1$ and for a multicomponent mixture of gases.

To express the metric in P coordinates, the following relationships are first written

$$\Delta s = \left[\frac{\partial s}{\partial T}\right] \Delta T + \sum_{i=1}^{r} \left[\frac{\partial s}{\partial \mu^i}\right] \Delta \mu^i, \quad \Delta a^i = \left[\frac{\partial a^i}{\partial T}\right] \Delta T + \sum_{j=1}^{r} \left[\frac{\partial a^i}{\partial \mu^j}\right] \Delta \mu^i \tag{8.53a, b}$$

Table 8.1 Thermodynamic potentials and the square of the element of thermodynamics length for four coordinate systems (Ruppeiner 1995)

Coordinates	Thermodynamic potential	$(\Delta l)^2$
$a = (u, \rho^1, \cdots, \rho^r)$	s	$-\frac{1}{k_B}\frac{\partial^2 s}{\partial a^\mu \partial a^\nu}\Delta a^\mu \Delta a^\nu$
$F = \left(\frac{1}{T}, -\frac{\mu^1}{T}, \ldots, -\frac{\mu^r}{T}\right)$	$\phi(F) = s - F^\mu a^\mu$	$\frac{1}{k_B}\frac{\partial^2 \phi}{\partial F^\mu \partial F^\nu}\Delta F^\mu \Delta F^\nu$
$P = (T, \mu^1, \ldots, \mu^r)$	$\omega(P) = u - Ts - \sum_{i=1}^{r} \mu^i a^i$	$-\frac{1}{k_B T}\frac{\partial^2 \omega}{\partial P^\mu \partial P^\nu}\Delta P^\mu \Delta P^\nu$
$a = (T, \rho^1, \ldots, \rho^r)$	$f = u - Ts$	$\frac{1}{k_B T}\left[\frac{\partial s}{\partial T}\right](\Delta T)^2 + \frac{1}{k_B T}\sum_{i,j=1}^{r} \frac{\partial \mu^i}{\partial a^j}\Delta a^i \Delta a^j$

Substituting these relationships into (8.45), one obtains

$$(\Delta l)^2 = -\frac{1}{k_B}\frac{\partial^2 \omega}{\partial P^\mu \partial P^\nu}\Delta P^\mu \Delta P^\nu \tag{8.54}$$

where

$$\omega(P^0, P^1, \ldots, P^r) = u - Ts - \sum_{i=1}^{r} \mu^i a^i \tag{8.55}$$

It may be easily shown that

$$\omega = -p = -\phi T \tag{8.56}$$

Table 8.1 shows the forms of the thermodynamic metric in various coordinate systems. There, ρ^i is the density of component i (the number of particles per unit volume).

8.3.2 The Analogy Fluid System—Magnetic System

A class of physical systems that allow an approach which is similar to the fluid systems is the class of magnetic systems. In terms of thermodynamics, the analogy between the two classes of systems is simple and will be presented for the case $n = 2$. Consider a system consisting of N magnetic spins placed in a magnetic field of intensity h. Denote by M the magnetization of the system, which is the sum of the magnetic moments of all spins in the system. It can be shown that the differential of the internal energy is given by:

$$dU = TdS + hdM \tag{8.57}$$

Here, the field h (but not the magnet generating this field) is assumed as making part of the system. On the other hand, for a fluid system of fixed volume V, it can be written

$$dU = TdS + \mu dN \tag{8.58}$$

Therefore, the following formal analogy between the two types of systems can be established:

$$N \leftrightarrow V, \quad M \leftrightarrow N, \quad m \leftrightarrow \rho \tag{8.59}$$

where $m \equiv M/N$ is the magnetization per spin.

This analogy have some theoretical difficulties, which finally can be resolved (see Ruppeiner 1995).

8.4 Covariant and Consistent Theory of Fluctuations

In this section it is presented a covariant theory of thermodynamic fluctuations. The main early contributions to this topic were made by Ruppeiner (1995). The starting point is the assumption that the probability density of the thermodynamic fluctuations is the solution of a set of partial differential equations of second order, linear parabolic type, whose coefficients are determined from the condition that in thermodynamic limit the solution converges towards the classical solution. This approach is similar in spirit to the theory of quantum mechanics, which starts from the assumption that the main tools are partial differential equations, which must meet a series of general principles.

8.4.1 *The Equation of the Probability Density of Fluctuations*

It is assumed that the probability density of fluctuations satisfies a generalized partial differential equation of diffusive type:

$$\frac{\partial P}{\partial t} = c_0(x)P + c_1^{\mu}(x)\frac{\partial P}{\partial x^{\mu}} + \frac{1}{2}c_2^{\mu\nu}(x)\frac{\partial^2 P}{\partial x^{\mu}\partial x^{\nu}} \tag{8.60}$$

In this equation, the "time" is not a measure of the real time, but it is a measure of the volume:

$$t = V^{-1} \tag{8.61}$$

The equation should be linear in P, in order to allow a normalization constant. The form of the Eq. (8.61) does not allow an explicit normalization. The normalization appears as an implicit relationship between the coefficients. A form of the equation that keeps the normalization is

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x^{\mu}}[K^{\mu}(x)P] + \frac{1}{2}\frac{\partial^2}{\partial x^{\mu}\partial x^{\nu}}[g^{\mu\nu}(x)P] \tag{8.62}$$

This is the equation which will be further used to describe the volume dependence of the thermodynamic fluctuations. It is formally identical to a Fokker–Planck equation and appears to be the simplest, but still general, mathematical formulation compatible with the theory of the thermodynamic fluctuations.

In Eq. (8.62), K^{α} is called *drift vector* and $g^{\alpha\beta}$ is the inverse of the metric tensor, which is supposed to be symmetric for the exchange of indices. Standard terminology is somehow improper because the quantity K^{α} does not actually transform itself like a contravariant tensor of the first rank. Generally, all coefficients appearing in (8.62) are functions of t and x. The absence of the dependence on t (i.e., indirect dependence on volume) constitutes the so-called hypothesis of translational invariance.

8.4.1.1 Average Values

Equation (8.62) can be used for the calculation of average values, which are defined as follows

$$<f> \; \equiv \int fPdx \quad (dx \equiv dx^0 dx^1 \cdots dx^r) \tag{8.63a, b}$$

Multiplication of (8.62) with the differential of the coordinates, followed by integration by parts, leads to:

$$\begin{aligned} &\frac{d}{dt}\int Pdx = 0, \quad \frac{d}{dt}<x^{\alpha}> \; = \; <K^{\alpha}>, \\ &\frac{d}{dt}<x^{\alpha}x^{\beta}> \; = \; <x^{\alpha}K^{\beta}> + <x^{\beta}K^{\alpha}> + <g^{\alpha\beta}> \end{aligned} \tag{8.64a–c}$$

The boundary contributions have not been considered here.

A specific case is considered now, i.e. the classical theory in the thermodynamic limit $t \to 0$. A Dirac delta function is adopted now as initial condition:

$$P_{t\to 0}(x|x_0) = \delta(x - x_0) \tag{8.65}$$

where x_0 represents the state of the infinite reservoir. From (8.64b) and (8.64c) it is found that

$$\lim_{t\to 0}\frac{d}{dt}<(x^{\alpha} - x_0^{\alpha})(x^{\beta} - x_0^{\beta})> \; = g^{\alpha\beta}(x_0) \tag{8.66}$$

or, in the first order in t:

$$< \Delta x^{\alpha} \Delta x^{\beta} > \; = g^{\alpha\beta}(x_0)t \tag{8.67}$$

This relationship is identical to (8.33). The obvious notation $\Delta x^{\alpha} = x^{\alpha} - x_0^{\alpha}$ was used here.

It can be shown that the mean values of the standard density coordinates are independent of t. For small values of t, the average value of $K^{\alpha}(x)$ is very close to its value assessed in x_0. Equation (8.64b) yields:

$$K^{\alpha}(a) = 0 \tag{8.68}$$

This result, together with the transformation rules for $K^{\alpha}(x)$ and $g^{\alpha\beta}(x)$, which will be presented below, uniquely determines the equation coefficients, for any coordinate system.

8.4.1.2 Thermodynamic Markovicity

Several essential properties arising from Eq. (8.62) will be presented in the following. First, note that a parabolic partial differential equation has a unique solution in a particular domain, if an initial condition for P is provided at some initial'time' t_0 and a specific boundary condition is also provided.

From a mathematical point of view, the process by which the initial condition at $t = t_0$ has been reached is not of particular importance. Such a statement is called the *Markov rule* and it motivates the usual notation of the density of probability:

$$P\left[\begin{array}{c|c} x & x_0 \\ t & t_0 \end{array}\right] \tag{8.69}$$

The Dirac delta function is in most cases used as an initial condition, because in the limit $t \to t_0$ the system A_V is in the same state as the system A_{V_0}. The Markov rule has an obvious physical significance: for given subsystem, at a given time (in a particular thermodynamic state), the probability of a fluctuation inside that subsystem depends only on the state of the subsystem and does not depend on the state of the environment.

The Markov rule is only an approximation, which is more suitable for short-range intermolecular potentials and for not very small subsystems. There are arguments pro and against the assumption that the validity of the Markov rule does not stop near the critical point (Ruppeiner 1995).

The Markov rule is implicit in the canonical ensemble, where equal probability is allocated to all microstates associated with given macroscopic parameters. However, the usage of the canonical ensemble is appropriate only in the thermodynamic limit. It is hoped that a weaker version of the Markov rule is still a good approximation in case of very small volumes.

8.4.1.3 The Chapman-Kolmogorov Equation

The solutions of the Eq. (8.62) are also solutions of the following Chapman-Kolmogorov equation

$$P\left[\begin{array}{c|c} x & x_0 \\ t & t_0 \end{array}\right] = \int P\left[\begin{array}{c|c} x & x_1 \\ t & t_1 \end{array}\right] P\left[\begin{array}{c|c} x_1 & x_0 \\ t_1 & t_0 \end{array}\right] dx_1 \tag{8.70}$$

with $t > t_1 > t_0$. The proof is done in two stages. Firstly, one checks that the right-hand side of Eq. (8.70) is a solution of Eq. (8.62), if

$$P\left[\begin{array}{c|c} x & x_1 \\ t & t_1 \end{array}\right] \tag{8.71}$$

is a solution. Secondly, the usage of the initial condition, which is a Dirac delta function, yields:

$$\lim_{t\to t_1, x=x_1} \int P\left[\begin{array}{c|c} x & x_1 \\ t & t_1 \end{array}\right] P\left[\begin{array}{c|c} x_1 & x_0 \\ t_1 & t_0 \end{array}\right] dx_1 = P\left[\begin{array}{c|c} x_1 & x_0 \\ t_1 & t_0 \end{array}\right] \tag{8.72}$$

that is just the required initial condition. The uniqueness theorem of the solutions of the Eq. (8.62), which is not presented here, completes the proof.

The classical theory of the thermodynamic fluctuations is not compatible with the Chapman-Kolmogorov equation, while the Gaussian approximation theory does satisfy that equation, as long as the metric elements do not depend on the thermodynamic state.

8.4.1.4 Covariance

The covariance of the partial differential equations of the thermodynamic fluctuations is discussed to some extent. First, the meaning of the covariance to a coordinate transformation from a set of thermodynamic variables x into another set $x'(x)$ should be explained. This is equivalent to saying that, when developing the new theory, by using the new coordinates, the new equation must be formally similar to the Eq. (8.62). Also, there is an explicit relationship between the probability density expressed in the new coordinates and the relationship expressed in the old coordinates, namely:

$$Pdx = P'dx' \tag{8.73}$$

This equation comes from the necessity that the probability of finding the particular state in a certain range of thermodynamic states does not depend on the coordinates used to specify that particular range. Finally, it must be demonstrated that the equation of the thermodynamic fluctuations in the new coordinates is equivalent with the fluctuations equation in the old coordinates. The fact that an equation of the form (8.62) satisfies all these requirements is not a priori so obvious. This is shown next.

In x' coordinates one can write:

$$\frac{\partial P'}{\partial t} = -\frac{\partial}{\partial x'^{\mu}}[K'^{\mu}P'] + \frac{1}{2}\frac{\partial^2}{\partial x'^{\mu}\partial x'^{\nu}}[g'^{\mu\nu}P'] \tag{8.74}$$

Next, the Eqs. (8.62) and (8.74) are systematically compared. This will show how to express K'^{α} and $g'^{\alpha\beta}$ in terms of K^{α} and $g^{\alpha\beta}$. First, one can write the equation:

$$P' = \left|\frac{dx}{dx'}\right| P \tag{8.75}$$

where the coefficient of P in the right-hand side is the Jacobian of the transformation of variables. Then, replace (8.75) in (8.74) and express the derivatives with respect to x', by using:

$$\frac{\partial}{\partial x'^{\alpha}} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial}{\partial x^{\mu}} \tag{8.76}$$

Compare with Eq. (8.62) by differentiating, dividing by the Jacobian and taking into account the equality of the corresponding coefficients of the derivatives of P. The following identities are used during calculation:

$$\begin{gathered}\frac{\partial}{\partial x^{\alpha}}\left|\frac{\partial x}{\partial x'}\right| = -\left|\frac{\partial x}{\partial x'}\right|\frac{\partial x^{\nu}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\mu}}{\partial x^{\alpha}\partial x^{\nu}}, \quad \frac{\partial}{\partial x^{\alpha}}\left|\frac{\partial x^{\beta}}{\partial x'^{\gamma}}\right| = -\frac{\partial x^{\mu}}{\partial x'^{\gamma}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial^2 x'^{\nu}}{\partial x^{\alpha}\partial x^{\mu}} \\ \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x'^{\mu}}{\partial x^{\beta}} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}\frac{\partial x^{\mu}}{\partial x'^{\beta}} = \delta^{\alpha}_{\beta}\end{gathered} \tag{8.77–79}$$

where the delta function of Kronecker has been used, which is defined as follows:

$$\delta^{\alpha}_{\beta} = \begin{cases} 0 & \text{if} \quad \alpha \neq \beta \\ 1 & \text{if} \quad \alpha = \beta \end{cases} \tag{8.80}$$

All derivatives with respect to x'^{α} can be placed in the Jacobian, or in the factors of

$$\frac{\partial x^{\alpha}}{\partial x'^{\beta}} \tag{8.81}$$

From the equality of the n^2 second order derivatives of P, it is obtained:

$$g'^{\alpha\beta} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} g^{\mu\nu} \tag{8.82}$$

From the equality of the n first derivatives of P, one finds:

$$K'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}} K^{\mu} + \frac{1}{2} g^{\mu\nu} \frac{\partial^2 x'^{\alpha}}{\partial x^{\mu} \partial x^{\nu}} \tag{8.83}$$

All the coefficients of the equation are now determined. Also, the factors containing only P are equals. This concludes the calculation required to prove the covariance of Eq. (8.62).

8.4.1.5 Translational Invariance

The theory is implicit in notation: neither K^{α} nor $g^{\alpha\beta}$ depend explicitly on t. This is the assumption of translational invariance, which was already mentioned. In the classical theory, this hypothesis is contained in the Axiom 8.1. The assumption is reasonable for volumes larger than the correlation volume, but it is questionable at smaller volumes.

8.4.1.6 Summary

The covariant theory of thermodynamic fluctuations presented here was developed based on five principles, which are listed below:

1. The average standard densities do not depend on volume;
2. The validity of a Chapman-Kolmogorov equation;
3. The theory is covariant;
4. The theory is consistent with the classical theory of fluctuations, in the limit of very large volumes;
5. The translational invariance.

Equation (8.62) satisfies all these five principles. Note that even the Gaussian approximation of the classical theory of fluctuations satisfies the five principles. However, this approximation is limited only to the case of very large volumes. When the volumes are not very large, the complete classical theory (beyond the Gaussian approximation) does not satisfy the first three principles.

8.4.2 *Fluctuations at Large Volumes*

The solutions of the equation of thermodynamic fluctuations can be naturally separated into three categories, corresponding to large volumes, intermediate volumes and low volumes, respectively. First, consider the case of large volumes, for which t is small. The probability distribution is strongly centered on the initial state x_0 and the coefficients K^α and the metric elements $g^{\alpha\beta}$ do not change very much in the range of states accessible to the system. As a first approximation, the coefficients are evaluated in x_0 and one can write:

$$\frac{\partial P}{\partial t} = -K^\mu \frac{\partial P}{\partial x^\mu} + \frac{1}{2} g^{\mu\nu} \frac{\partial^2 P}{\partial x^\mu \partial x^\nu} \tag{8.84}$$

Assume the initial condition is a Dirac delta function. Then, one can easily check that the normalized solution is:

$$P\left[\begin{array}{c|c} x & x_0 \\ t & 0 \end{array}\right] = \left[\frac{1}{2\pi t}\right]^{n/2} g^{1/2} \exp\left[-\frac{1}{2t} g_{\mu\nu}(\Delta x^\mu - K^\mu t)(\Delta x^\nu - K^\nu t)\right] \tag{8.85}$$

For any given value Δx^α, with t becoming smaller and smaller, this solution approaches the Gaussian approximation solution of the equation of the classical theory (8.29):

$$P\left[\begin{array}{c|c} x & x_0 \\ t & 0 \end{array}\right] = \left[\frac{1}{2\pi t}\right]^{n/2} g^{1/2} \exp\left[-\frac{1}{2t} g_{\mu\nu} \Delta x^\mu \Delta x^\nu\right] \tag{8.86}$$

The approximation solution of the Eq. (8.84) neglects the fact that the first derivatives of $g^{\alpha\beta}$ are, formally, of similar importance as the coefficients K^α. To see this, compute explicitly the derivatives in Eq. (8.62), before considering the coefficients as being constant:

$$\frac{\partial P}{\partial t} = \left[\frac{1}{2} g^{\mu\nu}_{,\mu\nu} - K^\mu_{,\mu}\right] P + (g^{\mu\nu}_{,\nu} - K^\mu) \frac{\partial P}{\partial x^\mu} + \frac{1}{2} g^{\mu\nu} \frac{\partial^2 P}{\partial x^\mu \partial x^\nu} \tag{8.87}$$

Next, the coefficients are evaluated in x_0. The solution is:

$$\begin{aligned} P\left[\begin{array}{c|c} x & x_0 \\ t & 0 \end{array}\right] &= c \left[\frac{1}{2\pi t}\right]^{n/2} g^{1/2} \exp\left[\left(\frac{1}{2} g^{\mu\nu}_{,\mu\nu} - K^\mu_{,\mu}\right) t\right] \\ &\times \exp\left[-\frac{1}{2} g_{\mu\nu} \left(\Delta x^\mu + g^{\mu\xi}_{,\xi} t - K^\mu t)(\Delta x^\nu + g^{\nu\xi}_{,\xi} t - K^\nu t\right)\right] \end{aligned} \tag{8.88}$$

where c is a constant. In the thermodynamic limit $t \rightarrow 0$, this solution is reduced to the Gaussian approximation of the classical theory. However, this solution does not preserve the normalization, because of the first exponential term.

8.5 Examples

Several examples of covariant calculus of fluctuations are presented in this section, for known thermodynamic systems (Ruppeiner 1995).

8.5.1 Paramagnetic System

Consider a subsystem A_N of N spins that do not interact each other, in an environment containing an infinite number of spins, characterized by temperature T and the external magnetic field intensity h. Each spin has two possible microstates, up and down, denoted for the spin i by $\sigma_i = +1$ or -1, respectively. Specifying all values σ_i is equivalent with the specification of subsystem's microstate.

The probability of finding a particular microstate of the system is given by a relationship known from statistical mechanics:

$$\frac{1}{Z}\exp\left(-\frac{H}{T}\right) \tag{8.89}$$

where the Hamiltonian of the system is

$$H = -h\sum_{i=1}^{N}\sigma_i \tag{8.90}$$

and the partition function is:

$$Z = 2^N \cosh^N\left(\frac{h}{T}\right) \tag{8.91}$$

A unit system for which $k_B = 1$ is adopted in this example and in the next one. The free energy per spin is given by:

$$\omega(T,h) = -\frac{T}{N}\ln Z = -T\ln\left|2\cosh\left(\frac{h}{T}\right)\right| \tag{8.92}$$

Knowing this relationship allows deducting all other thermodynamic relations. For example, the magnetization for the environment $A_{N_0}(N_0 \to \infty)$ is

$$m_0 = -\left[\frac{\partial \omega}{\partial h}\right]_T = \tanh\left(\frac{h}{T}\right) \tag{8.93}$$

and the entropy per spin, written as a function of magnetization, is

$$s(m) = -\left[\frac{\partial \omega}{\partial T}\right]_h = \ln\left[2\left(1/(1-m^2)\right)^{1/2}\right] - (\tanh^{-1} m)m \tag{8.94}$$

Since the entropy can be written as function of only one independent variable (i.e. the magnetization), the geometry of this system is one-dimensional.

The metric element comes from the two-times differentiation of $s(m)$ (see Table 8.1 and Eq. (8.59)):

$$g = \frac{1}{1-m^2} \tag{8.95}$$

The drift term K is null in these coordinates. Therefore, using Eq. (8.62) one finds the equation of the fluctuations

$$\frac{\partial P}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial m^2}\left[(1-m^2)P\right] \tag{8.96}$$

where $t = 1/N$. The equation is obeying the initial condition

$$P_{t\to 0}(m|m_0) = \delta(m - m_0) \tag{8.97}$$

and the boundary conditions for $m = \pm 1$, which preserve normalization

$$\frac{\partial}{\partial m}(g^{-1}P) = 0 \tag{8.98}$$

On the boundary, g^{-1} is zero. Therefore, the value $<m>$ is constant, as expected for a standard extensive parameter. Since the derivative of g^{-1} does not vanish on the boundary, the Eq. (8.98) requires the boundary condition $P = 0$.

The results obtained by using the covariant theory are closer to those obtained by using statistical mechanics, than the results predicted by the rigorous classical theory of fluctuations (i.e., without adopting the Gaussian approximation) (Ruppeiner 1995). Note that the covariant theory preserves the mean value $<m>$, while the rigorous classical theory does not preserve it.

8.5.2 The One-Dimensional Ising Model

The one-dimensional Ising model consists of a set of spins, as in Sect. 8.5.1, but this time they are arranged on a single line. In addition, each spin can interact with its closest neighbors, as shown by the Hamiltonian:

$$H = -J \sum_{i=1}^{\infty} \sigma_i \sigma_{i+1} - h \sum_{i=1}^{\infty} \sigma_i \tag{8.99}$$

where J is a coupling constant. In the ferromagnetic case, the adjacent spins tend to align with each other and the constant J is positive. In the antiferromagnetic case, the adjacent spins do not align each other and J is negative. Using the methods of statistical mechanics, one can show that the thermodynamic potential of this system is:

$$\phi(x, y) = \ln\left[e^x \cosh y + \left(e^{2x} \sinh^2 y + e^{-2x}\right)^{1/2}\right] \tag{8.100}$$

where $x \equiv J/T$ and $y \equiv -h/T$. The metric elements can be calculated easily in the coordinates F by using the thermodynamic length element found in the Eq. (8.47) and the Eq. (8.59) of the analogy magnetic system—fluid.

Results reported by Ruppeiner (1995) show that the covariant theory is more accurate than the rigorous classic theory (i.e., without adopting Gaussian approximation) even at very large volumes.

References

Ruppeiner, G.: Thermodynamics: a Riemanian geometric model. Phys. Rev. A **20**, 1608 (1979)
Ruppeiner, G.: Riemannian geometric theory of critical phenomena. Phys. Rev. A **44**, 3583 (1991)
Ruppeiner, G.: Riemannian geometry in thermodynamic fluctuation theory. Rev. Mod. Phys. **67**, 605–659 (1995)

Chapter 9

Thermodynamic Curvature. Correlation. Stability

The Gaussian approximation (8.86) for the fluctuations distribution loses its validity at large values of the parameter t. This has two explanations. First, the fluctuations become so large that the metric elements can not be considered as having constant values (evaluated at the initial state x_0). The second reason is that the drift terms proportional with t in the argument of the exponential inside the Eq. (8.85) become too large to be neglected.

For any value of t and any initial state x_0 it is possible to find a system of coordinates so that one or both of the above reasons become obvious. The limit of t for which the Gaussian approximation is valid in all coordinate systems is zero. The main problem is whether there is a limit of t so that the classical theory of the thermodynamic fluctuations ceases its validity in all systems of coordinates, for all values of t exceeding that limit.

At first sight it seems that the conversion rules (8.82) and (8.83) can always be used to find a coordinate system such that the drift vector K^{α} is globally null and for which the elements of the metric $g_{\alpha\beta}$ are constant, at global level. However, this is not possible, as it will be shown below.

Two ways to building an ideal coordinate system are quite obvious. The first method begins with a transformation towards a class of coordinate systems in which $K^{\alpha} = 0$, followed by a transformation to make the elements of the metrics $g_{\alpha\beta}$ constant at global level. The second method reverses the order of the two operations.

The first method may not lead to the desired results. Indeed, assuming that in the first stage it has been found a coordinate system for which $K^{\alpha} = 0$, this equality can be preserved through a linear transformation. But such a transformation, using Eq. (8.82), will not lead in general to constant elements $g_{\alpha\beta}$, unless the case when the transformation of the first stage has already made them constant. Another possible way is to choose certain nonlinear transformations which cancel the second term of the right-hand side of Eq. (8.83) and at the same time, transform the elements of the metric into constants. However, in general it is not possible to

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0_9

obtain the second effect, and even less possible to obtain both effects, as it will be shown below.

The reasoning outlined above leads to the conclusion that, for a given thermodynamic system, the existence of an ideal system of coordinates it is almost ruled out. But this way of thinking does not lead to finding the limiting value of t, i.e. the value for which the classical theory should cease to be valid, for any system of coordinates. For this reason, the second procedure of the first attempt to obtain constant metric elements should be studied in more details.

9.1 Equivalent Metrics

In geometry, the following problem may be formulated. For a given metric $g_{\alpha\beta}(x)$, can a new coordinate system be found such that the metric has constant elements throughout the space? To a similar, but more general question, Riemann provided an answer in 1861. Next, some details on this issue are given, following the excellent approach by Ruppeiner (1995).

The question is whether, given a coordinate system x, with the square of the length of a line element given by:

$$dl^2 = g_{\mu\nu}(x)dx^{\mu}dx^{\nu} \tag{9.1}$$

it is possible to find another coordinate system $x' = x'(x)$, having a *given* metric $g'_{\alpha\beta}(x')$. The restriction is that the coordinate transformation $x'(x)$ keeps the distance between any pair of neighboring points, i.e.:

$$g_{\mu\nu}(x)dx^{\mu}dx^{\nu} = g'_{\mu\nu}(x')dx'^{\mu}dx'^{\nu} \tag{9.2}$$

Riemann showed that this is possible, if and only if, a certain quantity is transformed as a fourth rank tensor.

A metric relationship of the type (9.2) can be expressed as a partial differential equation of the first order for $x'(x)$:

$$g_{\alpha\beta}(x) = \frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}g'_{\mu\nu}(x') \tag{9.3}$$

Both metrics are symmetric, so that only $n(n+1)/2$ of their elements are independent. It is noted that in (9.3) there are $n(n+1)/2$ independent equations, with only n functions. It is clear that, for these n functions to constitute a solution, they have to satisfy $n(n+1)/2 - n$ additional relationships.

The idea to be followed is as follows. It is assumed that the solution exists, and the conditions of existence of that solution should be found. For this, it is considered a point x_0 and one writes:

$$x'^{\alpha}(x) = x_0'^{\alpha} + x'^{\alpha}_{,\mu}\Delta x^{\mu} + \frac{1}{2!}x'^{\alpha}_{,\mu\nu}\,\Delta x^{\mu}\Delta x^{\nu} + \frac{1}{3!}x'^{\alpha}_{,\mu\nu\xi}\,\Delta x^{\mu}\Delta x^{\nu}\Delta x^{\xi} + \cdots \tag{9.4}$$

where the coefficients $\Delta x^{\alpha} = x^{\alpha} - x_0^{\alpha}$ and all coefficients of the derivative-type are evaluated in x_0. The derivatives can be computed by repeated differentiation of the Eq. (9.3) and solving relatively simple algebraic equations. But the procedure for obtaining the derivatives has a potential difficulty. Indeed, the coefficients of the series that are obtained might have not all of their lower indices commutable. Such invariance of the change of indices is essential, because the order of the partial differentiation should be of no importance.

First, the focus is on the first derivative and calculations are performed. When passing to the next derivative, only the change of the two last indices will be considered, as the other indices already commute, due to processing of the lower-order derivatives. The first group of algebraic equations consists of the partial differential Eq. (9.3), evaluated in the point x_0:

$$g_{\alpha\beta} = x'^{\mu}_{,\alpha}x'^{\nu}_{,\beta}g'_{\mu\nu} \tag{9.5}$$

It is said that g and g' are congruent matrices in point x_0. It can be shown that if one matrix is symmetric, then the other matrix must be symmetric, too.

Of fundamental importance for building a local Cartesian coordinate system is the theorem which states that any positive definite symmetric matrix is congruent to the identity matrix.

By differentiation of (9.3) with respect to x_{γ} one finds:

$$\begin{aligned} g_{\alpha\beta,\gamma} = {} & \frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}\frac{\partial x'^{\xi}}{\partial x^{\gamma}}g'_{\mu\nu,\xi} + \frac{\partial x'^{\mu}}{\partial x^{\beta}}\frac{\partial^2 x'^{\nu}}{\partial x^{\alpha}\partial x^{\gamma}}g'_{\mu\nu} \\ & + \frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial^2 x'^{\nu}}{\partial x^{\beta}\partial x^{\gamma}}g'_{\mu\nu} \end{aligned} \tag{9.6}$$

where the derivatives of $g_{\mu\nu}$ and $g'_{\mu\nu}$ are carried out in relation to their natural variables x and x'.

To find the second derivatives, $x'^{\alpha}_{,\beta\gamma}$, it is useful to introduce the Christoffel symbols of second kind (see Eq. (4.149)):

$$\Gamma^{\alpha}_{\beta\gamma} \equiv \frac{1}{2}g^{\mu\alpha}\left(g_{\mu\beta,\gamma} + g_{\mu\gamma,\beta} - g_{\beta\gamma,\mu}\right) \tag{9.7}$$

It is immediately apparent that the Christoffel symbols are symmetrical in the two lower indices:

$$\Gamma^{\alpha}_{\beta\gamma} = \Gamma^{\alpha}_{\gamma\beta} \tag{9.8}$$

A similar definition exists for the coordinates x'. After computation with indexed quantities and using the relationship:

$$g'^{\alpha\beta} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} g^{\mu\nu} \tag{9.9}$$

one finds:

$$\frac{\partial^2 x'^{\alpha}}{\partial x^{\beta} \partial x^{\gamma}} = \Gamma^{\mu}_{\beta\gamma} \frac{\partial x'^{\alpha}}{\partial x^{\mu}} - \Gamma'^{\alpha}_{\mu\nu} \frac{\partial x'^{\mu}}{\partial x^{\gamma}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \tag{9.10}$$

Since the Christoffel symbols are symmetric, the right-hand side of the equation is invariant with respect to the interchange of β and γ, as wanted. By evaluation in the point x_0, the second order coefficients of the Taylor series are found. Equation (9.10) is differentiated with respect to x^{δ} and the following relationship is found:

$$\begin{aligned} \frac{\partial^3 x'^{\alpha}}{\partial x^{\beta} \partial x^{\gamma} \partial x^{\delta}} = & - \Gamma^{\mu}_{\beta\delta} \Gamma'^{\alpha}_{\nu\xi} \frac{\partial x'^{\nu}}{\partial x^{\gamma}} \frac{\partial x'^{\xi}}{\partial x^{\mu}} - \Gamma^{\mu}_{\beta\gamma} \Gamma'^{\alpha}_{,\nu\xi} \frac{\partial x'^{\nu}}{\partial x^{\delta}} \frac{\partial x'^{\xi}}{\partial x^{\mu}} \\ & - \Gamma^{\mu}_{\gamma\delta} \Gamma'^{\alpha}_{\nu\xi} \frac{\partial x'^{\nu}}{\partial x^{\mu}} + \Gamma'^{\mu}_{\nu\xi} \Gamma'^{\alpha}_{\mu o} \frac{\partial x'^{\nu}}{\partial x^{\delta}} \frac{\partial x'^{\xi}}{\partial x^{\gamma}} \frac{\partial x'^{o}}{\partial x^{\beta}} + \Gamma'^{\mu}_{\nu\xi} \Gamma'^{\alpha}_{\mu o} \frac{\partial x'^{\nu}}{\partial x^{\delta}} \frac{\partial x'^{\xi}}{\partial x^{\beta}} \frac{\partial x'^{o}}{\partial x^{\gamma}} \\ & + \Gamma^{\mu}_{\beta\gamma} \Gamma^{\nu}_{\mu\delta} \frac{\partial x'^{\alpha}}{\partial x^{\nu}} + \Gamma^{\mu}_{\beta\gamma,\delta} \frac{\partial x'^{\alpha}}{\partial x^{\mu}} - \Gamma'^{\alpha}_{\mu\nu,\xi} \frac{\partial x'^{\mu}}{\partial x^{\gamma}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \frac{\partial x'^{\xi}}{\partial x^{\delta}} \end{aligned} \tag{9.11}$$

where the second order derivatives were eliminated by replacing the known values (9.10). Third order derivatives of (9.11) are already symmetric to the interchange of β cu γ, but they are not necessarily symmetric to the interchange of γ with δ. It is requires that:

$$x'^{\alpha}_{,\beta\gamma\delta} - x'^{\alpha}_{,\beta\delta\gamma} = 0 \tag{9.12}$$

Performing the calculations shows that this condition is fulfilled if and only if the following quantity with four indices

$$R^{\alpha}_{\beta\gamma\delta} = \Gamma^{\alpha}_{\beta\gamma,\delta} - \Gamma^{\alpha}_{\beta\delta,\gamma} + \Gamma^{\mu}_{\beta\gamma} \Gamma^{\alpha}_{\mu\delta} - \Gamma^{\mu}_{\beta\delta} \Gamma^{\alpha}_{\mu\gamma} \tag{9.13}$$

(with a similar definition in the coordinates x') is transformed as a tensor of rank four, namely:

$$R^{\alpha}_{\beta\gamma\delta} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \cdot \frac{\partial x'^{\nu}}{\partial x^{\beta}} \cdot \frac{\partial x'^{\xi}}{\partial x^{\gamma}} \cdot \frac{\partial x'^{o}}{\partial x^{\delta}} \cdot R'^{\mu}_{\nu\xi o} \tag{9.14}$$

$R^{\alpha}_{\beta\gamma\delta}$ is called the *Riemann (curvature) tensor* and its transformation rule is precisely the necessary relation looked for.

The proof is virtually complete, since the fourth order coefficients $x'^{\alpha}_{,\beta\gamma\delta\varepsilon}$ should have the last two indices commutable, both δ and ε coming out from the differentiating of a single expression. This is maintained for higher order coefficients, too.

Equation (9.14) is a necessary condition for the existence of a solution of the partial differential Eq. (9.3). It allows building a solution in terms of Taylor series. Riemann showed that the relationship (9.14) is the sufficient condition, for the particular case of zero curvature. Later, Christoffel demonstrated that relationship (9.14) is a sufficient condition for the existence of the solution in the general case.

9.2 Properties of Riemann Curvature Tensor

Several properties that are easy to demonstrate of the Riemann tensor are shown below. From (9.13) one finds:

$$\begin{aligned} R^{\alpha}_{\beta\gamma\delta} &= -R^{\alpha}_{\beta\delta\gamma} \\ R^{\alpha}_{\beta\gamma\delta} + R^{\alpha}_{\gamma\delta\beta} + R^{\alpha}_{\delta\beta\gamma} &= 0 \end{aligned} \tag{9.15, 16}$$

The following tensor with four covariant indices, obtained by using the coefficients of the fundamental form, is important:

$$R_{\alpha\beta\gamma\delta} = g_{\alpha\mu} \cdot R^{\mu}_{\beta\gamma\delta} \tag{9.17}$$

Calculation using (9.7), (9.13) and the next identities, quite easy to prove:

$$\begin{aligned} g_{\alpha\mu} \cdot g^{\mu\beta}_{,\gamma} &= -g_{\alpha\mu,\gamma} \cdot g^{\mu\beta} \\ g_{\alpha\beta,\gamma} &= g_{\alpha\mu} \cdot \Gamma^{\mu}_{\beta\gamma} + g_{\beta\mu} \cdot \Gamma^{\mu}_{\alpha\gamma} \end{aligned} \tag{9.18, 19}$$

demonstrates that:

$$\begin{aligned} R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}, \quad R_{\alpha\beta\gamma\delta} &= -R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\delta\gamma} = R_{\beta\alpha\delta\gamma} \\ R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} &= 0 \end{aligned} \tag{9.20–22}$$

Define the Ricci tensor of second rank:

$$R_{\alpha\beta} = R^{\mu}_{\alpha\mu\beta} \tag{9.23}$$

It can be rewritten as:

$$R_{\alpha\beta} = g^{\mu\nu} \cdot R_{\mu\alpha\nu\beta} \tag{9.24}$$

Equation (9.20) leads to the conclusion that the Ricci tensor is symmetric. The Riemann curvature scalar is defined as follows:

$$R = g^{\mu\nu} \cdot R_{\mu\nu} \tag{9.25}$$

The value of R in any point is independent of the choice of coordinates. Therefore R is a scalar.

Note that other notations and conventions of sign are sometimes used in literature for the previous relationships (as seen in previous chapters).

9.3 Normal Riemann Coordinates

Consider again the problem of finding a coordinate system for which the metric elements are constant. In such a coordinate system, all curvature components should be zero, as follows from the definition (9.13) of the curvature. But, if the curvature is zero in one coordinate system, it must be zero in all coordinate systems; this is a result of the tensor transformation (9.14). Therefore, the cancellation of the Riemann tensor is a necessary condition for the existence of a coordinate system whose metric components are constant. This is equally well a sufficient condition, because canceling the Riemann curvature allows satisfying the consistency Eq. (9.14) with a constant metric.

All these results are valid only locally. There is no consequence on the global topology. For example, the Riemann curvature of a cylinder is zero, but the cylinder is not topologically equivalent to a plan. The thermodynamic curvature of a perfect gas is null, as shown below. However, in the general case of non-ideal gases, the curvature is not zero.

The following question arises: how far from the point x_0 the state of the system can be moved, but still keeping the system geometry reasonably close to that of a flat surface? To answer this question, the first problem is to find a coordinate system $x'(x)$ as close as possible to the Cartesian coordinates near the point x_0. Without loss of generality, the assumption $x'_0 = 0$ is adopted here.

The first condition is that the elements of the metric are (locally) Cartesian in x'_0:

$$g'_{\alpha\beta} = \delta_{\alpha\beta} \tag{9.26}$$

This is done easily by solving the algebraic Eq. (9.5) in the first derivatives $x'^{\alpha}_{,\beta}$. The number of unknowns is larger than the number of equations and, consequently, the additional degrees of freedom enable a change of orientation of the coordinate system.

Next, try to cancel in x_0 as many derivatives of $g'_{\alpha\beta}(x)$ as possible. Start with the first derivatives and take:

$$\Gamma'^{\alpha}_{\beta\gamma} = 0 \tag{9.27}$$

Using (9.27) and (9.10) yields the second derivatives coefficients:

$$x'^{\alpha}_{,\beta\gamma} = \Gamma^{\mu}_{\beta\gamma} \cdot x'^{\alpha}_{,\mu} \tag{9.28}$$

However, the second derivatives of $g'_{\alpha\beta}(x)$ cannot all be zero, because this would mean a zero curvature, which is not wanted. This is the end of the process of canceling the derivatives of the metric elements.

One can make improvements by looking to the geodesics. The first step is to observe that, since the first derivatives of $g'_{\alpha\beta}$ can be canceled, it follows that:

$$R'^{\alpha}_{\beta\gamma\delta} = \Gamma'^{\alpha}_{\beta\gamma,\delta} - \Gamma'^{\alpha}_{\beta\delta,\gamma} \tag{9.29}$$

where the left hand side is already known from the transformation rule of the curvature tensor, Eq. (9.14). The first derivatives of the Christoffel symbols may be chosen in such a way that this consistency relationship is satisfied. However, the choice is not unique. Indeed, one can add a symmetric quantity in indices γ and δ to the first derivatives of the Christoffel symbols without changing the difference in Eq. (9.29). The specific choice:

$$\Gamma'^{\alpha}_{\beta\gamma,\delta} = \frac{1}{3}\left(R'^{\alpha}_{\beta\gamma\delta} + R'^{\alpha}_{\gamma\beta\delta}\right) \tag{9.30}$$

fulfils the Eq. (9.29), as it can be easily checked by using the Eqs. (9.15) and (9.16). Then, substituting (9.30) in (9.11), would get $x^{\alpha}_{,\beta\gamma\delta}$. These coordinates are named *normal Riemann coordinates*. It can be shown that they are the closest coordinates to the local Cartesian coordinates. The coefficients of the type of higher derivatives are not unique.

In the following the prime sign is dropped since the original coordinate system will not be used. A central notion in Riemann geometry is the geodesic, which is the path of the shortest distance between points. The length of a path $x(\lambda)$ is given by the following integral equation:

$$\int \left(\left| g_{\mu\nu} \frac{dx^{\mu}}{d\lambda} \frac{dx^{\nu}}{d\lambda} \right| \right)^{1/2} d\lambda \tag{9.31}$$

where λ is the arc length of the curve. It has been shown in previous chapters that the geodesics are solutions of the geodesics equation, which, written in terms of the parameter λ, have the form:

$$\frac{d^2x^{\alpha}}{d\lambda^2} + \Gamma^{\alpha}_{\mu\nu} \frac{dx^{\mu}}{d\lambda} \frac{dx^{\nu}}{d\lambda} = 0 \tag{9.32}$$

Without loss of generality, one can choose $x_0 = 0$, and λ can be measured in relation with this point. The geodesic curves passing through $x_0 = 0$ can be expressed by using the expansion in Taylor series:

$$x^\alpha = b^\alpha \lambda + \frac{1}{2!} c^\alpha \lambda^2 + \frac{1}{3!} d^\alpha \lambda^3 + \cdots \tag{9.33}$$

The development of the Christoffel symbols in a Taylor series, keeping the first terms, is:

$$\Gamma^\alpha_{\mu\nu}(x) = \Gamma^\alpha_{\mu\nu} + \Gamma^\alpha_{\mu\nu,\xi} b^\xi \lambda + \cdots \tag{9.34}$$

Substituting this series in the geodesics equation written in Riemann normal coordinates, and equating the similar powers of λ, one finds that the term with c^α vanishes. Indeed, the Christoffel symbols are canceled and d^α cancels due to the antisymmetry of the Riemann tensor in its last two indices. The result is:

$$x^\alpha = b^\alpha \lambda + \cdots \tag{9.35}$$

In a curved space, this equation is the closest to that of a flat space.

9.4 The Validity of the Classical Theory of Fluctuations

Recall that the objective is to assess how much the system state can be moved with respect to the point x_0 so that its geometry remains almost flat. A two-dimensional (2D) geometry is now used, where a measure of the deviation with respect to flatness is the circumference of a circle of (small) radius r, centered in $x_0 = 0$. The circle is the geometric locus of those points on the geodesics that start from x_0, with parameter $\lambda = r$.

One develops in Taylor series the metric elements:

$$g_{\alpha\beta} = \delta_{\alpha\beta} + \frac{1}{2} g_{\alpha\beta,\mu\nu} b^\mu b^\nu \lambda^2 + \cdots \tag{9.36}$$

The direction cosines are express parametrically:

$$b^1 = \cos\theta, \quad b^2 = \sin\theta \tag{9.37}$$

where θ varies from 0 to 2π. In normal Riemann coordinates, Eq. (9.19) leads to:

$$g_{\alpha\beta,\gamma\delta} = \frac{1}{3}\left(R_{\alpha\gamma\beta\delta} + R_{\beta\gamma\alpha\delta}\right) \tag{9.38}$$

The relationship that gives the scalar Riemann curvature in terms of the Riemann tensor is also used:

$$R = \frac{2R_{1212}}{g} \tag{9.39}$$

which is valid in any coordinate system. The length of the circumference of the circle is given by the integral:

$$C = \int_{\theta=0}^{2\pi} \left(|g_{\mu\nu} \frac{dx^{\mu}}{d\theta} \cdot \frac{dx^{\nu}}{d\theta}| \right)^{1/2} \cdot d\theta \tag{9.40}$$

Replacing the expansion of the metric and the parametric expressions of b^{α}, and performing laborious calculations, one finally obtains:

$$C = 2\pi r + \frac{\pi}{6} R r^3 + \cdots \tag{9.41}$$

In the Riemann geometry of thermodynamics, the square of the distance has dimension of inverse volume. In conjunction with (9.41), one sees that R must have the dimension of volume. For the geometry to be almost flat, the first term of the right member of (9.41) must dominate the second term. This condition is obtained if:

$$r^2 \ll \frac{12}{|R|} \tag{9.42}$$

It can be shown (Ruppeiner 1995) that typical fluctuations act on a distance of the order of

$$r^2 \approx \frac{2}{V} \tag{9.43}$$

Therefore, in order for the classical theory to give a correct description of the fluctuations, the following condition should be fulfilled:

$$V \gg \frac{1}{6} |R| \tag{9.44}$$

So, $|R|$ fixes the lower limit of the volume for the validity of the classical theory.

9.5 Weinhold Geometry

In a series of papers, Weinhold (1975, 1976) proposed a geometrical representation of the standard equilibrium thermodynamic. Weinhold used the energy representation, denoting the $n+1$ extensive variables of the subsystem by $Y_e = (S, N^1, \ldots, V, \ldots, N^n)$. These variables are similar to those used in the entropy representation, with the difference that in the first position appears the entropy instead of the internal energy. The last variable sets the scale of the subsystem; it can be the volume. The conjugated intensive variables are:

$$P_e^\alpha = \frac{\partial U}{\partial Y_e^\alpha} \tag{9.45}$$

i.e. $P_e = (T, \mu^1, \ldots, -P, \ldots, \mu^r)$. Taking into account the Gibbs–Duhem relationship, any of the quantities P_e^α can be expressed as a function of the other quantities. The dependent variable P_e^α is usually listed at the end, being the conjugate of the fixed scale of the subsystem.

The differentials of the thermodynamic functions can be associated with vectors in a vector space. The basis of the vector space may be built by using the n independent values dP_e^α, since:

$$df = \sum_{\mu=0}^{n-1} \frac{\partial f}{\partial P_e^\mu} dP_e^\mu \tag{9.46}$$

for any thermodynamic function f. Weinhold associated the inner product:

$$\left(dP_e^\alpha \,|\, dP_e^\beta \right) = \frac{\partial^2 U}{\partial Y_e^\alpha \partial Y_e^\beta} \tag{9.47}$$

to the pairs of the basis vectors. The inner product is commutative. Also, Weinhold imposed the condition that the inner product is bilinear:

$$\left(dP_e^\alpha \,|\, a dP_e^\beta + b dP_e^\gamma \right) = a\left(dP_e^\alpha \,|\, dP_e^\beta \right) + b\left(dP_e^\alpha \,|\, dP_e^\gamma \right) \tag{9.48}$$

for any constant values of a and b.

In the thermodynamic space of phases one can define, in principle, any kind of inner product. To make the inner product useful in applications, it must have a physical motivation. Weinhold's idea was to associate the inner product, positive defined, with the second law of thermodynamics. It can be shown that:

$$\sum_{\mu,\nu=0}^{n-1} \frac{\partial^2 U}{\partial Y_e^\mu \partial Y_e^\nu} dY_e^\mu dY_e^\nu = -T \sum_{\mu,\nu=0}^{n-1} \frac{\partial^2 S}{\partial X_e^\mu \partial X_e^\nu} dX_e^\mu dX_e^\nu \tag{9.49}$$

The right hand side of (9.49) is a positive definite quadratic form. This is enough to prove that the matrix of (9.47) is positive definite. Thus:

$$\left(dP_e^\alpha \mid dP_e^\alpha\right) \geq 0 \tag{9.50}$$

and

$$\left(dP_e^\alpha \mid dP_e^\alpha\right)\left(dP_e^\beta \mid dP_e^\beta\right) - \left(dP_e^\alpha \mid dP_e^\beta\right)^2 \geq 0 \tag{9.51}$$

These inequalities are ways of expressing the second law of thermodynamics.

Weinhold geometry can be developed by using the entropy representation, because the sign of the inequalities does not change when the inner product is multiplied by any positive function. The entropy representation is usually preferred, due to its applications in the theory of fluctuations and in the theory of finite time thermodynamics.

The inner product introduced by Weinhold may lead to a Riemann metric, as shown in (9.49). However, such geometry does not seem to have physical interpretation in the strict context of equilibrium thermodynamics. Meanings can be found in the context of fluctuations theory or of finite time processes theory.

9.6 Thermodynamic Curvature

For a given thermodynamics state, the value of the scalar curvature is independent of the coordinate system in which calculations are conducted. Therefore, the choice of the coordinates is arbitrary. This observation is essential because it endows the curvature with intrinsic physical significance.

9.6.1 Simple Examples

In the following there are presented some examples of simple systems for which the calculation of the thermodynamic curvature may be conducted in detail. Also, the results are discussed and interpreted (for more details see Ruppeiner 1995).

9.6.1.1 One-Component Perfect Gas

The perfect gas model is well-known. The molecules are considered identical and do not interact with one another except when they are in direct contact.

Calculations can be done by starting either from the energy representation $U = U(S, V)$ or from the entropy representation $S = S(U, V)$. The way these representations can be obtained in case of a monoatomic perfect gas is briefly presented. It can start from simple expressions, such as that of the internal energy and that of the state equation, respectively:

$$U = \frac{3}{2} NRT, \quad PV = NRT \tag{9.52, 53}$$

where N is the amount of substance (in mole) and R is the universal constant of perfect gases. The well-known relation which expresses the internal energy in terms of the pressure P and volume V is inferred from the relationships (9.52) and (9.53):

$$U = \frac{3}{2} PV \tag{9.54}$$

Next, consider the expression of the perfect gas entropy in terms of volume and pressure:

$$S = S_0 + (R + C_V) \ln \frac{V}{V_0} + C_V \ln \frac{P}{P_0} \tag{9.55}$$

Here C_V is the molar heat at constant volume (which, in the case of the perfect gas, is a constant) and the quantities with subscript 0 represent integration constants. From Eq. (9.55) the pressure can be expressed under the form $P = P(S, V)$. If this relationship is replaced in (9.54), a relationship is obtained in which the only variables are U, S and V. By extracting the internal energy U from that relation, the explicit energy representation $U = U(S, V)$ is found. Similarly, by extracting the entropy S, the explicit representation of entropy $S = S(U, V)$ is obtained. The thermodynamic curvature calculation can be performed by starting from any of the two representations. However, a different approach, proposed by Ruppeiner (1995), is presented here.

The Helmholtz free energy per unit volume, f, is defined by the standard relationship

$$f = u - Ts \tag{9.56}$$

where u and s are the internal energy and the entropy, respectively, both of them per unit volume. For the perfect gas, the explicit relationship which gives the Helmholtz free energy is:

$$f(T, \rho) = \rho k_B T \ln \rho + \rho k_B f_1(T) \tag{9.57}$$

where $\rho = N/V$ is the gas density, expressed in number of moles per unit volume, k_B is Boltzmann constant and $f_1(T)$ is a function of temperature (having negative second derivative, in order to yield to a positively definite heat capacity). The entropy s and the chemical potential μ may be obtained by using the equations:

$$s = -\left(\frac{\partial f}{\partial T}\right)_{\rho}, \quad \mu = \left(\frac{\partial f}{\partial \rho}\right)_{T} \tag{9.58, 59}$$

It can be shown that, by using these relationships, one obtains the following diagonalized expression of the square of the thermodynamic length Δl (Ruppeiner 1995):

$$\Delta l^2 = -\frac{\rho \cdot f_1''(T)}{T}(\Delta T)^2 + \frac{1}{\rho}(\Delta\rho)^2 \tag{9.60}$$

where the sign $''$ designates the second derivative. It is useful to show that the thermodynamic curvature is canceled by a change to the Euclidean coordinates. First, define:

$$t = \int\limits_{T_0}^{T} \left(-\frac{f_1''(T)}{T}\right)^{1/2} dT \tag{9.61}$$

where T_0 is a positive arbitrary constant. Then, define the following two coordinates:

$$x^1 = (2\rho)^{1/2}\left[\cos\left(\frac{t}{2}\right) + \sin\left(\frac{t}{2}\right)\right], \quad x^2 = (2\rho)^{1/2}\left[\cos\left(\frac{t}{2}\right) - \sin\left(\frac{t}{2}\right)\right] \tag{9.62, 63}$$

Direct calculations lead to an expression of the square of the thermodynamic length of the form:

$$\Delta l^2 = \left(\Delta x^1\right)^2 + \left(\Delta x^2\right)^2 \tag{9.64}$$

which refers to the length of a locally Euclidean line segment. This observation is enough to understand that the Riemann curvature of the perfect gas is zero. It can be shown that the curvature of a perfect gas whose scale is fixed by the number of particles instead of its volume, is also null. The conclusion that the thermodynamics curvature of the perfect gas is zero is an important result, which obvious connects the thermodynamics curvature with the interaction between particles.

Also, a null curvature is obtained from the original metric (9.60), by using the equation of the Riemann curvature. For a diagonal metric,

$$\Delta l^2 = g_{11}\left(\Delta x^1\right)^2 + g_{22}\left(\Delta x^2\right)^2 \tag{9.65}$$

one can easily find from the Eq. (9.39) that the scalar Riemann curvature is given by:

$$R = \frac{1}{g^{1/2}}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{g^{1/2}}\frac{\partial g_{22}}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{g^{1/2}}\frac{\partial g_{11}}{\partial x^2}\right)\right] \tag{9.66}$$

The replacement of the line element (9.60) yields to a zero curvature.

If the metric is not diagonal, but has the form:

$$\Delta l^2 = g_{11}\left(\Delta x^1\right)^2 + 2g_{12}\,\Delta x^1 \Delta x^2 + g_{22}\left(\Delta x^2\right)^2 \tag{9.67}$$

it is easy to prove that the expression of the scalar Riemann curvature is (Ruppeiner 1995):

$$R = -\frac{1}{g^{1/2}}\left[\begin{array}{l} \dfrac{\partial}{\partial x^1}\left(\dfrac{1}{g^{1/2}}\dfrac{g_{12}}{g_{11}}\cdot\dfrac{\partial g_{11}}{\partial x^2} - \dfrac{1}{g^{1/2}}\cdot\dfrac{\partial g_{22}}{\partial x^1}\right) \\ \quad + \dfrac{\partial}{\partial x^2}\left(\dfrac{2}{g^{1/2}}\cdot\dfrac{\partial g_{12}}{\partial x^2} - \dfrac{1}{g^{1/2}}\cdot\dfrac{\partial g_{11}}{\partial x^2} - \dfrac{1}{g^{1/2}}\cdot\dfrac{g_{12}}{g_{11}}\cdot\dfrac{\partial g_{11}}{\partial x^1}\right) \end{array}\right] \tag{9.68}$$

Note that, even when the curvature is null, the topology of this surface endowed with Riemann geometry is not that of a flat surface.

9.6.1.2 Ideal Paramagnetic System

Another simple application refers to an ideal paramagnetic system consisting of magnetic spins without interaction. The system is characterized by the following equation of state:

$$m = f\left(\frac{h}{T}\right), \tag{9.69}$$

where m is the spin magnetization, h is the intensity of the magnetic field and f is a function with positive first derivative. The state equation yields the relationship:

$$\left(\frac{\partial u}{\partial h}\right)_T = 0 \tag{9.70}$$

The length of the line segment in (T, m) coordinates is, according to the analogy between a fluid and a magnetic system:

$$\Delta l^2 = g_{TT}\cdot(\Delta T)^2 + g_{mm}(\Delta m)^2 \tag{9.71}$$

where

$$g_{TT} = \frac{1}{k_B T}\left(\frac{\partial s}{\partial T}\right)_m, \quad g_{mm} = \frac{1}{k_B T}\left(\frac{\partial h}{\partial m}\right)_T \tag{9.72, 73}$$

Using the magnetic equation of state (9.69), the next relationship is found, which is useful to find the extreme:

$$\left(\frac{\partial g_{mm}}{\partial T}\right)_m = 0 \tag{9.74}$$

Also, Eq. (9.70) gives:

$$\left(\frac{\partial g_{TT}}{\partial m}\right)_T = 0 \tag{9.75}$$

Replacing into the thermodynamic curvature Eq. (9.66), one sees that the curvature of the ideal paramagnetic system is zero, provided that the following metric element does not cancel:

$$g_{TT} \neq 0 \tag{9.76}$$

If, however, this metric element is null, as in the case when the spins and the lattice do not have internal energy levels, the geometry is singular and actually one-dimensional. From the mathematical point of view, this also leads to a zero curvature, since in a one-dimensional geometry one can always find an Euclidean coordinate system.

9.6.1.3 Mixture of Perfect Gases

The Helmholtz free energy per unit volume of a mixture of r perfect gases is (Ruppeiner 1995):

$$f\left(T, a^1, a^2, \cdots, a^r\right) = \sum_{i=1}^{r} a^i k_B T \ln a^i + \sum_{i=1}^{r} a^i k_B f_i(T) \tag{9.77}$$

where a^i is the density of gas i (number of particles per unit volume) and the functions $f_i(T)$ depend on temperature and have negative second derivatives, in order to obtain a positive caloric capacity. Using the expression of the free energy in the metric equation, one can finally get the square of the length of a line segment under the following diagonal form:

$$\Delta l^2 = \left[\sum_{i=1}^{r} a^i \cdot h_i(T)\right](\Delta T)^2 + \sum_{i=1}^{r} \frac{1}{a^i} \cdot \left(\Delta a^i\right)^2 \tag{9.78}$$

where

$$h_i(T) = -\frac{f_i''(T)}{T} > 0 \tag{9.79}$$

Since the metric is simple, many of the elements of the curvature tensor are canceled. Using Eq. (9.25), the scalar curvature is obtained:

$$R = \frac{\sum_{i=1}^{r} h_i(T) \cdot \sum_{j=1}^{r} a^j \cdot h_j(T) - \sum_{i=1}^{r} a^i [h_i(T)]^2}{2\left[\sum_{i=1}^{r} a^i \cdot h_i(T)\right]^2} \tag{9.80}$$

If all functions $h_i(T)$ are identical, as in case when all the components of the mixture are monoatomic gases, the Eq. (9.80) reduces to:

$$R = \frac{r-1}{2\rho} \tag{9.81}$$

where

$$\rho = \sum_{i=1}^{r} a^i \tag{9.82}$$

is the total density of the mixture. This result is simple and surprising. It depends neither on temperature nor on the relative densities of the component gases but only on the total density of the gas mixture.

If the “mixture” consists of a single gas, $r = 1$ and this yields $R = 0$, which is to be expected.

9.6.1.4 Perfect Quantum Gases

It is possible to calculate the thermodynamic curvature of a Fermi gas, as well as of a Bose gas (Janyszek 1986; Ruppeiner 1995). The thermodynamic potential to be used is:

$$\Phi\left(\frac{1}{T}, -\frac{\mu}{T}\right) = k_B(2S+1)\lambda^{-3} f_{\pm}\left(\frac{5}{2}, \eta\right) \tag{9.83}$$

where

$$\lambda = \frac{h}{(2\pi m k_B T)^{1/2}} \tag{9.84}$$

is the thermal wavelength and the following two notations were used

$$\eta = e^{\frac{\mu}{k_B T}}, \quad f_\pm(l, \eta) = \frac{1}{\Gamma(l)} \int_0^\infty \frac{x^{l-1}}{\eta^{-1} e^x \pm 1} dx \tag{9.85, 86}$$

In these relationships h is Planck's constant, m is the mass of the particle, S is spin of the particle and Γ is the Euler gamma function. Also, a Fermi gas corresponds to a plus sign and a minus sign means a Bose gas.

In coordinates $\Gamma = (1/T, -\mu/T)$, the elements of the metric are:

$$g_{\alpha\beta} = \frac{1}{k_B} \cdot \Phi_{,\alpha\beta} \tag{9.87}$$

It can be shown that, if the metric elements can be written as second order derivatives of a specific potential, the thermodynamics curvature is given by (Ruppeiner 1995):

$$R = \frac{k_B}{2} \frac{\det\begin{pmatrix} \Phi_{,11} & \Phi_{,12} & \Phi_{,22} \\ \Phi_{,111} & \Phi_{,112} & \Phi_{,122} \\ \Phi_{,112} & \Phi_{,122} & \Phi_{,222} \end{pmatrix}}{\left[\det\begin{pmatrix} \Phi_{,11} & \Phi_{,12} \\ \Phi_{,12} & \Phi_{,22} \end{pmatrix}\right]^2} \tag{9.88}$$

This relationship is obtained from Eq. (9.39) of the thermodynamics curvature. Note that it contains second and third order derivatives. After calculating the derivatives, one finds:

$$R = -\frac{5\lambda^3}{(2S+1)} \frac{\left[\begin{array}{r} f_\pm^2\left(\frac{3}{2}, \eta\right) f_\pm\left(\frac{1}{2}, \eta\right) - 2 f_\pm\left(\frac{5}{2}, \eta\right) f_\pm^2\left(\frac{1}{2}, \eta\right) \\ + f_\pm\left(\frac{5}{2}, \eta\right) f_\pm\left(\frac{3}{2}, \eta\right) f_\pm\left(-\frac{1}{2}, \eta\right) \end{array}\right]}{\left[5 f_\pm\left(\frac{5}{2}, \eta\right) f_\pm\left(\frac{1}{2}, \eta\right) - 3 f_\pm^2\left(\frac{3}{2}, \eta\right)\right]^2} \tag{9.89}$$

For a Fermi gas, the curvature R is always positive while for a Bose gas, R is always negative and strongly diverge when $T \to 0$. This limit is in connection with Bose-Einstein condensation. In both cases (Bose and Fermi), R tends to zero in the limit of the perfect gas.

9.6.1.5 Van der Waals Gas

Details about the calculation of the curvature of a closed system consisting of a Van der Waals gas are presented in the following. In this case, the Helmholtz free energy per unit volume is

$$f(T,\rho) = \rho k_B T \ln\left(\frac{\rho}{\rho_0}\right) + \rho k_B e(T) - \rho k_B T \ln(1 - b\rho) - a k_B \rho^2 \tag{9.90}$$

where $e(T)$ is a function of the temperature, having negative second derivative, while a,b and ρ_0 are constants. The constant heat capacity per molecule is denoted c_v. A units system where c_v is dimensionless quantity of the order unity, for example 3/2 for a mono-atomic perfect gas, is considered here. Then,

$$e(T) = -c_v T \ln\left(\frac{T}{T_0}\right) + \varepsilon_0 \tag{9.91}$$

where T_0 and ε_0 are constants. The metric is diagonal, as seen from Eq. (8.52). The metric elements are:

$$g_{TT} = \frac{\rho c_v}{T^2}, \quad g_{\rho\rho} = \frac{1}{\rho(1 - b\rho)^2} - \frac{2a}{T} \tag{9.92}$$

The curvature is calculated by using Eq. (9.66):

$$R = \frac{-(1 - b\rho)\begin{bmatrix} -2a^2\rho + 6a^2 b\rho^2 - 6a^2 b^2 \rho^3 + 2a^2 b^3 \rho^4 + 2aT \\ -a c_v T - 2ab\rho T + 3abc_v \rho T - 3ab^2 c_v \rho^2 T \\ + ab^3 c_v \rho^3 T + bc_v T^2 \end{bmatrix}}{c_v(-2a\rho + 4ab\rho^2 - 2ab^2\rho^3 + T)^2} \tag{9.93}$$

Near the critical point, the metrics element $g_{\rho\rho}$ and its first derivative with respect to ρ are zero. The thermodynamics curvature near the critical point is computed as follows. The critical parameters of the system are defined by using standard relationships:

$$T_c = \frac{8a}{27b}, \quad \rho_c = \frac{1}{3b} \tag{9.94}$$

Using (9.93) and replacing the critical density ρ_c and the reduced temperature $\tau \equiv T/T_c$, one finds, in the first order of approximation in τ,

$$R = -\frac{3b}{2c_v}\frac{1}{\tau^2} \tag{9.95}$$

The quantity R is of the order of the molecular volume.

Very close to the critical point, the specific heat is given by

$$C_\rho = -T_c \frac{\partial^2 f}{\partial T^2} = \frac{c_v k_B}{3b} \tag{9.96}$$

and the combination

$$R C_\rho \tau^2 = -\frac{1}{2} k_B \tag{9.97}$$

is valid in first order of approximation in τ.

It can be shown that, in case of the van der Waals gas, the curvature does not have a particular behavior for first order phase transitions. From Eq. (9.93) it is seen that a divergence can occur only when the denominator is canceled, and this happens at the critical point.

9.6.2 *Thermodynamics Curvature and Correlation Length*

The thermodynamics curvature has volume units. It has been seen that the curvature is zero in case of the ideal gas, for which no long distance interactions between particles exist. The curvature tends to infinity at the critical point, in a similar way with the correlation volume ξ^d (here ξ is the correlation length and d is the spatial dimension). These properties suggest that thermodynamics curvature can be used as a measure of the degree of interaction between the particles of the system.

The theory also showed that the thermodynamics curvature can be used as a measure of the smallest volume for which the classical fluctuations theory, based on the assumption of a homogeneous medium, is valid. Near the critical point this volume is expected to be ξ^d. In other words, it is expected that a relationship of proportionality between the curvature and correlation volume does exist. Assume that this relation has the following form:

$$R = \kappa_2 \xi^d \tag{9.98}$$

where κ_2 is a dimensionless constant of the order of unity.

The thermodynamics curvature can be linked with another aspect of the theory of fluctuations, which is usually overlooked. First, the free energy per unit volume is used as a thermodynamic potential

$$\phi = s - \frac{1}{T}\mu + \frac{\mu}{T}\rho \tag{9.99}$$

Assume that the energy and the number of particles are conserved. Then, the deviation of the entropy with respect to its maximum value is

$$\Delta S_0 = V\left[\Delta s - \frac{1}{T_0}\Delta u + \frac{\mu_0}{T_0}\Delta\rho\right] \tag{9.100}$$

where the subscript 0 refers to the parameters of the very large system A_{V_0}.

The fluctuations of temperature and chemical potential tend to zero near the critical point. On the contrary, the fluctuations of the entropy and standard extensive quantities close to the critical point become very large. This can be shown by standard approaches. Therefore, the quantity in the brackets of the Eq. (9.100) is very close to the differential of ϕ and

$$\Delta S_0 \approx V\Delta\phi \tag{9.101}$$

For common size fluctuations the following relation applies:

$$|\Delta S_0| \approx k_B \tag{9.102}$$

Therefore

$$|\Delta\phi| \approx \frac{k_B}{V} \tag{9.103}$$

If the magnitude of the fluctuations of the thermodynamic potential per unit volume is in the same range with magnitude of the potential, one can write:

$$\Delta\phi \approx \phi \tag{9.104}$$

Equation (9.103), together with (9.104), leads to:

$$\phi = -\frac{\kappa_1 k_B}{\xi^d} \tag{9.105}$$

where κ_1 is a dimensionless constant of the order of unity. This is sometimes called the hyperscaling hypothesis. Note that ϕ refers only to the singular part of the free energy, i.e. that part associated with long-range interactions between parts of the system. But ξ diverges at the critical point. Consequently, ϕ must tend to zero at the critical point.

Therefore, ϕ may be determined, at least in the first order of approximation, by subtracting from the actual value of the free energy, the value of the free energy at the critical point.

One may use the Eq. (9.98) to eliminate the correlation volume from (9.105). The result is:

$$R = -\kappa \frac{k_B}{\phi}, \quad (\kappa \equiv \kappa_1 \kappa_2) \tag{9.106}$$

This equation states the proportionality between the thermodynamics curvature and the inverse of the free energy.

Next, an example will be shown by using a simple model.

9.6.2.1 One-Dimensional Ising Model

The thermodynamics of the one-dimensional Ising model is relatively simple. It can be shown that the following expression gives the thermodynamics curvature:

$$-R = \cosh y\left(\sinh^2 y + e^{-4x}\right)^{-1/2} + 1 \tag{9.107}$$

Here the notation of Sect. 8.5.2 is used. The correlation length was assessed using statistical mechanics in the ferromagnetic case ($J > 0$), where the correlation function is monotonously damped (Ruppeiner 1995). It was found that the value of the Gauss curvature:

$$\xi_G \equiv -\frac{R}{2} \tag{9.108}$$

is in very good agreement with the correlation length, without deviating from it by more than a single lattice constant. In the anti-ferromagnetic case $(J < 0)$, the dampening of the correlation function oscillates, with an wavelength of the order of the lattice parameter, while ξ_G is of the order of separation distance between the neighboring spins.

The results are compelling for the ferromagnetic case $J > 0$ but those for the anti-ferromagnetic case $J < 0$ are not self-evident. For a brief discussion, some statistical mechanics arguments are used. Consider the following new variables:

$$\sigma_i' = (-1)^i \sigma_i, \quad h_i' = (-1)^i h \tag{9.109a, b}$$

By using the transformation

$$J' = -J > 0 \tag{9.110}$$

the following Hamiltonian is obtained

$$H = -J' \sum_{i=1}^{\infty} \sigma_i' \sigma_{i+1}' - \sum_{i=1}^{\infty} h_i' \sigma_i' \tag{9.111}$$

which is identical with that obtained in the ferromagnetic case $J > 0$. As a result, by using the new variables (9.109), the same thermodynamics curvature and correlation length are obtained as in the case $J > 0$.

Without the use of variables (9.109), it can be shown that for $h = 0$ the spins are locked into a single microstate, with alternating spins on a length corresponding to the correlation length ξ. The correlation length ξ diverges when $T \to 0$. Therefore, ξ gives the lower limit for the volume where thermodynamics can be used. In this situation, thermodynamics curvatures of the order of magnitude of the lattice parameter do not have, of course, physical significance.

9.6.3 Thermodynamics Curvature and Stability

In this chapter it was argued that the thermodynamics curvature may be used as a measure for the correlation length. Janyszek and Mrugala (1989a, b) proposed another interpretation of the thermodynamics curvature. These authors suggested that R is a measure of stability. Thus, the higher R is (with the sign convention) the more stable the system is. This interpretation is motivated, in part, by the results of the one-dimensional Ising model, where the magnetization in the ferromagnetic case has large fluctuations at low temperature and zero magnetic field intensity, but the fluctuations in the anti-ferromagnetic case are small. Therefore, in the first case, large groups of spins align up or down, and the system is less stable.

Another motivation lies in the results obtained for ideal quantum gases of bosons and fermions, respectively. These systems, which are fundamentally different from each other, have opposite thermodynamic curvatures. For a Bose gas, the curvature R is negative and for a Fermi gas the curvature is positive, which shows that the Bose gas, with attractive interactions, is less stable.

The interpretation of the thermodynamic curvature as a measure of systems stability is appealing. However, it has some weaknesses, such as that it does not allow an interpretation of the meaning of the units of thermodynamic curvature and does not offer an explanation for the fact that the correlation length is of the order of magnitude of the thermodynamic curvature.

References

Janyszek, H.: On the Riemannian metric structure in the classical statistical equilibrium thermodynamics. Rep. Math. Phys. **24**(1), 1–10 (1986)

Janyszek, H., Mrugala, R.: Geometrical structure of the state space in classical statistical and phenomenological thermodynamics. Rep. Math. Phys. **27**(2), 145–159 (1989a)

Janyszek, H., Mrugala, R.: Riemannian geometry and the thermodynamics of model magnetic systems. Phys. Rev. A **39**(12), 6515–6523 (1989b)

Ruppeiner, G.: Riemannian geometry in thermodynamic fluctuation theory. Rev. Mod. Phys. **67**, 605–659 (1995)

Weinhold, F.: Metric geometry of equilibrium thermodynamics. J. Chem. Phys. **63**, 2479 (1975)

Weinhold, F.: Metric geometry of equilibrium thermodynamics. V. Aspects of heterogeneous equilibrium. J. Chem. Phys. **65**, 559 (1976)

Index

V. Badescu, *Modeling Thermodynamic Distance, Curvature and Fluctuations*,
Understanding Complex Systems, DOI 10.1007/978-3-319-33789-0