AN AXIOMATIZATION OF CLASSICAL PHENOMENOLOGICAL THERMODYNAMICS

W.J. HORNIX

AN AXIOMATIZATION OF CLASSICAL PHENOMENOLOGICAL THERMODYNAMICS

by

W. J. HORNIX

Department of Applied Mathematics and Mathematical Physics,
University College of South Wales and Monmouthshire, Cardiff.

A thesis submitted in candidature for the degree
of Philosophiae Doctor in the University of Wales.

July 1970.

Acknowledgments

I wish to express my sincere gratitude to Professor P. T. Landsberg for giving me the opportunity of working in his Department and for his encouragement and his participation in so many penetrating discussions which shaped this work to its present state.

I am indebted to Mr. D. A. Evans for correcting my English and for his continuous care in the period that this thesis was typed and I was abroad. Thanks also to Mrs. Hudson for typing the manuscript.

The work for this thesis was made possible through a Research Fellowship of the Royal Society of London held in 1967-68 in this Department, and through financial assistance of the Foundation "Clementer et Auguste" in Amsterdam for the period 1968-69. I wish to acknowledge this with gratitude.

CONTENTS

PREFACE

This thesis contains two parts. Chapters 1 - 6 contain an exposition of an axiomatisation of classical phenomenological thermodynamics of closed systems. The framework of this part will be published in the Proceedings of the International Symposium "A Critical review of the foundations of relativistic and classical thermodynamics", University of Pittsburgh (Pa) 1969 (1). The extension of this axiomatization to open systems is prepared in a contribution for the International Conference on Thermodynamics, Cardiff, 1970, and will be published in the Proceedings of this Conference (2). The basic ideas explained in the latter article need further development. Thermodynamics of open systems falls outside the scope of this thesis, except for a few remarks.

The proposed axiomatization is characterised as follows:

1 The set of states Z_i of a system is structured through

(a) the existence of "connections", defined as equivalence relations, generating the classes of equal temperature and generalised forces;

(b) the existence of "isolations", defined as equivalence relations generating the classes of equal entropy, internal energy and dimensions; the adiabatic isolation is a simple preorder in Z_i.

2 On the basis of accessibility relations for adiabatic isolation and for energetic isolation for the composition of systems we can define extensive entropy functions S_i and internal energy functions U_i. A similar procedure can lead to extensive deformation coordinate functions X_{ki}.

3 The maximum entropy principle and a minimum energy principle in a local formulation, and certain continuity assumptions imply the existence of absolute temperature and absolute force functions. By strengthening the extremal principles and the assumptions concerning the occupation of phase space, it is possible to describe the behaviour of systems in the domain of negative absolute temperatures and at the boundaries of phase space.

The original traits of this axiomatization, apart from improvements of certain parts of earlier work in the field, mentioned below, are:

The construction of an extensive internal energy function with the help of a system which can be interpreted as a calorimeter : the measurement of internal energy differences is thus reduced to caloric measurements.

The formulation of the maximum entropy principle as the fundamental relationship of the extensive and intensive variables of a thermodynamic system.

The introduction of a minimum internal energy principle which leads to the definition of absolute force functions and which establishes the independence of thermodynamics of mechanics.

The analysis of the properties of thermodynamic systems at the boundaries of phase space, including a restatement of the third law.

The second part is a critical analysis of other axiomatisations which cover the same field, and which are attempts to develop a theory, starting from accessibility relations. Within these limits fall the article of G. Falk and H. Jung in the Handbuch der Physik (3), R. Giles' Mathematical Foundations of Thermodynamics (4), J. L. B. Cooper's article in the Journal of Mathematical Analysis and Applications (5) and J. J. Duistermaat's article in Synthese (6). L. Tisza's contributions, collected in Generalised Thermodynamics (7) lie outside this frame. In many respects his work starts where my work ends, with a small overlap in Chapters 5 and 6 of this thesis. Similar remarks can be made with respect to P. T. Landsberg's Thermodynamics (8), which may be considered as a preparation of much work in the field of axiomatics of thermodynamics, through a careful analysis of the presuppositions of the traditional presentations of the subject.

The discussion of the contributions 3 - 6 is undertaken because there is certainly a lack of mutual criticism in the field of axiomatisation of thermodynamics. A criticism of the work of Falk and Jung does not exist and only Duistermaat gives a more than incidental comment on Giles' work. However, Duistermaat's approach is so closely related to Giles' that objections against the latter contribution hold also for the former. This critical analysis should not be considered as negative: the attention given to these books and articles is not least intended to express esteem for the valuable work in the field delivered by the authors.

REFERENCES

(1) W. J. Hornix, An axiomatization of classical phenomenological thermodynamics. To be published in the Proceedings of the International Symposium 'A critical review of the Foundations of Relativistic and Classical Thermodynamics,' University of Pittsburgh, Pa, 1969.

(2) W. J. Hornix, The extension of classical phenomenological thermodynamics to open systems. To be published in the Proceedings of the International Conference on Thermodynamics, Cardiff, 1970, in Pure and Applied Chemistry.

(3) G. Falk and H. Jung, Handbuch der Physik, Band III/2, Springer pp.119-175 Berlin, 1959.

(4) R. Giles, Mathematical Foundations of Thermodynamics, Pergamon, 1964.

(5) J. L. B. Cooper, J. Math. Anal. Appl. 17, 172 (1967).

(6) J. J. Duistermaat, Synthese 18, 311 (1968).
(7) L. Tisza, Generalised Thermodynamics, M.I.T. Press, 1966.
(8) P. T. Landsberg, Thermodynamics, Interscience, 1961.

CHAPTER I INTRODUCTION

The aim of this introductory chapter is to give an outline of the formal theory, as developed in detail in the following chapters, to give an interpretation of its primitive terms, and to make general remarks about its relationships to the traditional presentation of thermodynamics. Traditional thermodynamics includes different approaches: the Kelvin-Clausius approach, the theory developed by Carathéodory, and the Gibbsian approach. It is neither my intention to deny the differences in these three presentations, nor to assess these differences. I will refer to "traditional thermodynamics" only to clarify the physical relevance of the formal structure given.

The outline of the formal theory is given in the form of a set theoretical definition of the concept"set $\mathcal{Z}$ of closed, commensurable, thermodynamic systems". In such a definition the primitive terms of the formalised theory are enumerated as well as the primitive relationships applying to the theory. The axioms of the following chapters can thus be considered as a restatement for reasons of convenience at a place where we actually use those propositions. To avoid unnecessary repetitions I will not formulate all the axioms explicitly in the set theoretical definition of this chapter, but refer to the precise statements in the following chapters. Different parts of the definition will be followed by short indications about the results which follow within the formal theory from the foregoing statement. I do not intend to give precise "rules of interpretation" for the primitive terms. This would presuppose a sophisticated use of the language of experimental physics and of a "meta language" which connects experimental physics with our formal theory and neither of the two are sufficiently specified to fulfill this task. The remarks about the interpretation of the primitive terms are meant as an indication how the statements of the theory can be translated in physical terms, and they will make it possible to judge the relevance of the formal theory of thermodynamics as a physical science.

General remarks and notes with respect to the interpretation will succeed immediately those parts of the definition to which they refer. The formal theory can however be considered as self-contained and thus independent of these remarks. To distinguish the formal theory proper from other parts of the discourse I will separate these parts by horizontal lines. Formal parts start and end with the sign Δ.

Systems, states.

Δ

A set $\mathcal{Z}$ of closed, commensurable, thermodynamic systems is a set $\mathcal{Z}$ of sets (or formal systems) Z_i of elements (states) z_i such that:

1.1 if $Z_i, Z_j \in \mathcal{Z}$ then $Z_i \times Z_j \in \mathcal{Z}$.

Δ

The set $Z_i \equiv \{z_i, z_i', z_i'' \text{---}\}$ is interpreted as the set of all possible equilibrium states of a closed physical system, say G_i, the latter being defined through a description in terms of experimental physics. The interpretation of the cartesian product of the sets Z_i and Z_j, or the "composition" of the formal systems Z_i and Z_j, is that it represents the set of all equilibrium states of the system, consisting of the physical systems G_i and G_j, combined physically in such a way, that every duplet $z_i\, z_j$ can be considered as an equilibrium state of the combination, and as the conjunction of the states z_i and z_j of G_i and G_j separately: i.e. the systems G_i and G_j are "not connected at all" i.e. "isolated with respect to each other."

Connections, thermal and generalized force connections

Δ

1.2 There exists a "direct thermal connection", C_θ, being an equivalence relation on $\bigcup_{i,j} Z_i \times Z_j$; $Z_i, Z_j \in \mathcal{Z}_\theta \subset \mathcal{Z}$.

1.3 If $Z_i, Z_j \in \mathcal{Z}_\theta$, then $(Z_i \times Z_j) \cap C_\theta \in \mathcal{Z}_\theta$.

Δ

A connection C for the family $\mathcal{Z}$ of sets (formal systems) Z_i, Z_j, --- can be defined generally as a relation on the set $\bigcup_{i,j} Z_i \times Z_j$, this being the set of all duplets $z_i z_j$ which can be formed from the elements (states) of the sets $Z_i \in \mathcal{Z}$ (also included the duplets of identical states $z_i z_i$). A relation on $\bigcup_{i,j} Z_i \times Z_j$ is a subset of this set. A connection C can be interpreted as the set of equilibrium states $z_i z_j$ of a physical combination of the physical systems G_i and G_j (with sets of possible equilibrium states Z_i and Z_j respectively) such that not every duplet of equilibrium states of G_i and G_j is also an equilibrium state of the combined system. Thus we call two systems connected if not every pair of equilibrium states of the two systems separately is an equilibrium state of the combination.

A special class of connections are those connections which have the formal structure of an equivalence relation.

An equivalence relation on $\bigcup_{i,j} Z_i \times Z_j$ is a relation ϱ such that for all $z_i \in Z_i$, $z_j \in Z_j$ and $z_k \in Z_k$

(i) $z_i z_i \in \varrho$ "reflexivity"

(ii) if $z_i z_j \in \varrho$ then $z_j z_i \in \varrho$ "symmetry"

(iii) if $z_i z_j \in \varrho$ and $z_j z_k \in \varrho$ then $z_i z_k \in \varrho$ "transitivity"

It will be immediately clear that many ways of connecting physical systems in practical experiment have the properties associated with equivalence relations. Usually these properties are considered as self evident and not worthwhile mentioning. Only the transitive property of the thermal connection gets a different treatment in traditional thermodynamics.

The set $\mathcal{Z}_\theta$ can be interpreted as the set of closed systems for which a unique temperature function is defined, i.e. systems which have a temperature, and which are not divided by adiabatic walls which make it possible to have different temperatures in different parts.

The set $(Z_i \times Z_j) \cap C_\theta$ contains all duplets $z_i z_j$ of equal temperature and is thus the set of equilibrium states of the thermally connected systems Z_i and Z_j; sometimes we write for $Z_i \times Z_j \cap C_\theta$: $Z_i(\theta) Z_j$.

The statement of the existence of a thermal connection is closely related to the introduction of an empirical temperature function via the zeroth law. The zeroth law formulates the transitive property of the thermal connection. The difference between our approach and the zeroth law approach is that we do not use the much stronger assumption that the thermal connection relation, restricted to the domain $Z_i \times Z_j$ is an analytical function of a complete set of independent variables of the two systems, which would necessitate us to introduce deformation - and force-coordinates, and that we are precise in the formulation of the mathematical properties of the thermal connection relation. The existence of a real valued empirical temperature function will consequently not be proved, and appears superfluous for the derivation of the existence of an absolute temperature function with the desired thermodynamic properties. The thermal connection relation decides only whether or not two states are in thermal equilibrium (say "have equal temperatures"), an order is not established.

Δ

1.4 There exist "direct generalized force connections", C_{ϕ_k}, $k \in N$, where N is a set of integers indicating "different kinds" of force connections; C_{ϕ_k} being equivalence relations on $\bigcup_{i,j} Z_i \times Z_j$; $Z_i, Z_j \in \mathcal{Z}_{\phi_k} \subset \mathcal{Z}$.

1.5 If $Z_i, Z_j \in \mathcal{Z}_{\phi_k}$, then $(Z_i \times Z_j) \cap C_{\phi_k} \in \mathcal{Z}_{\phi_k}$.

Δ

In traditional presentations of thermodynamics force-variables are tacitly introduced as acceptable coordinates describing the state of a system. The philosophy behind it is presumably that a foreknowledge of mechanics, electrostatics, etc. is presupposed and that only the thermodynamic concepts proper: temperature, heat, internal energy and entropy need definition, and that these definitions can be given purely in terms of the presupposed disciplines (e.g. the "mechanical definition of heat"). The philosophy behind our approach is quite different: it seems impossible to reduce thermodynamics to mechanics etc; there are necessarily non-mechanical primitive terms in thermodynamics, in

this axiomatization "thermal connection", "adiabatic isolation" and "energetic isolation". That the latter concept is thermodynamical will be clear if we remind that it leads to the definition of an internal energy function and internal energy cannot be identified with mechanical energy in case of non mechanical systems. It seems worthwhile trying to formulate thermodynamics without foreknowledge of other parts of physics, and it is perhaps possible to consider mechanics (or at least statics) as a special case of thermodynamics (thermostatics), namely "thermodynamics" restricted to those systems, which do not contain thermal systems and for which all pairs of states are reversible adiabatically accessible. This last remark needs further clarification, which however cannot be given here.
An attempt to present thermodynamics without foreknowledge of mechanics forces us to introduce "mechanical" concepts on the same level as thermodynamic concepts. The force concept can be introduced in striking similarity with the temperature concept. In a first stage we state the existence of generalized force connections as equivalence relations: such a relation does nothing more than reply "yes" or "no" to the question whether two states of a certain class of systems are in k^{th} force equilibrium. A real valued force function is not defined through it. In a second stage we can define an "absolute force" which can be interpreted as the force variable which is traditionally borrowed from mechanics etc.
For the description of all the possible ways in which human ingenuity can connect systems and which could be called "force connection" a perhaps unlimited list of equivalence relations has to be drafted. For instance, two quantities of a fluid can be connected by means of a movable piston (pressure connection) or enclosed in rigid containers each provided with a movable piston, which pistons are connected through a shaft or a spring. (Mechanical force connections of a different kind). Chemical cells can be

connected with conducting wires (electromotoric force connection) etc.

We restrain from specifications by introducing the concept of "generalized forces" and restrict ourselves to those connections which can be described as equivalence relations ("direct" connections). The sets $\mathcal{Z}_{\phi_k}$ can be interpreted as the class of closed systems for which the connection C_{ϕ_k} is defined, i.e. for which a k^{th} force coordinate exists or can be connected in a way, k, specified in experimental physics, and for which the k^{th} force coordinate is uniquevalued.

The set $(Z_i \times Z_j) \cap C_{\phi_k}$ (or $Z_i(\phi_k) Z_j$) has to be interpreted as the set of equilibrium states of the systems Z_i and Z_j, connected with respect to the k^{th} force coordinate.

Isolations, adiabatic, energetic and dimensional isolations

Δ

1.6 There exist for all $Z_i \in \mathcal{Z}$ "adiabatic isolations" $(Is)_i$ or $\Rightarrow_i$, being simple preorder relations defined on $Z_i \times Z_i$; $Z_i \in \mathcal{Z}$, and such that "extensivity or metrization axioms for entropy" hold (axiom 2.9)

1.7 There exists an entropymeter Z_σ (definition 2.17) and all $Z_i \in \mathcal{Z}$ are S-measurable (definition 2.21) with respect to this meter (axiom 2.22)

This leads to the definition and existence proof of a set of extensive or additive entropyfunctions $S_i(z_i)$ for the domains Z_i, $Z_i \in \mathcal{Z}$.

Δ

In traditional thermodynamics isolations are usually defined as specified relationships between a system and its environment. Many times these definitions take the form of negative propositions; e.g. adiabatic isolation is defined as the non existence of thermal connections between the system and its environment. This invokes the difficulty of the introduction of the environment of a system and the inconvenience of definitions in the form of negative statements. In case of adiabatic isolation we can get around the last by a definition of the form "a system is adiabatically isolated, if all the connections between the system and its environment are force connections". But this policy compels us to list all force connections, also the less obvious "indirect" ones, and does not relieve us of the task of defining the environment of a system.

These difficulties do not arise if we define isolations through "accessibility relations", which explain which processes are possible under the given isolation (in case of the adiabatic isolation) or which states are linked by possible processes (in case of the other isolations). The adiabatic isolation has the structure of a simple preorder relation.

A simple preorder on the domain Z_i is a relation ρ on $Z_i \times Z_i$ such that, for all $z_i, z_i', z_i'', z_i''' \in Z_i$,

(i) $\langle z_i z_i \rangle \in \rho$ reflexivity

(ii) $\langle z_i' z_i'' \rangle \in \rho$ or $\langle z_i'' z_i' \rangle \in \rho$ (or both) comparability

(iii) if $\langle z_i' z_i'' \rangle \in \rho$ and $\langle z_i'' z_i''' \rangle \in \rho$ then $\langle z_i' z_i''' \rangle \in \rho$ transitivity

The duplet of states $\langle z_i' z_i'' \rangle$ which belong to the adiabatic isolation relation $(\mathrm{Is})_i$ must be interpreted as the initial state and final state of a possible adiabatic process. A much simpler notation will be the replacement of $\langle z_i' \; z_i'' \rangle \; \varepsilon \; (Is)_i$ by $z_i' \Rightarrow_i z_i''$, or more shortly $z_i' \Rightarrow z_i''$. The omission of the index i is allowed because the indices of the states before and after the arrow determine already the system for which adiabatic isolation is meant.

The axioms for the simple preorder can thus be rewritten for the case of adiabatic isolation as follows:

For all $z_i, z_i', z_i'', z_i''' \; \varepsilon \; Z_i$

(i) $z_i \Rightarrow z_i$ reflexivity

(ii) $z_i' \Rightarrow z_i''$ or $z_i'' \Rightarrow z_i'$ (or both) comparability

(iii) if $z_i' \Rightarrow z_i''$ and $z_i'' \Rightarrow z_i'''$ then $z_i' \Rightarrow z_i'''$ transitivity

This notation has the additional advantage that it is immediately clear that $z_i' \Rightarrow z_i''$ is the representation of a possible process. We will reserve the notation $\langle z_i' \; z_i'' \rangle$ to indicate a more general "process", impossible or possible. This latter duplet of states must be well distinguished from the duplet $z_i' \; z_i''$ which indicates an equilibrium state of the composition $Z_i \times Z_i$.
The metrization axioms explain how the adiabatic isolations $\Rightarrow_{ijk \ldots}$ of compositions of systems $Z_i \times Z_j \times Z_k \times$ --- are related to those of the component systems $\Rightarrow_i, \Rightarrow_j, \Rightarrow_k$. They aim to be sufficient and necessary conditions (together with the proposition stating the existence of a special system, the entropy meter) for the existence of entropy functions with additive properties. After this stage this approach comes together with that in which the existence of such an additive entropy function is assumed axiomatically.
(Tisza (7), Callen H. B. Thermodynamics (Wiley N.Y. 1960))
The term "possible adiabatic process" needs specification in the light of the formal properties of the adiabatic isolation relation. In the first place: processes which can be approximated as

closely as desired by possible processes will also be called "possible": thus adiabatically reversible processes are possible in both directions.

In the second place, the fact that an adiabatic isolation is defined also for the composition $Z_i \times Z_j$ of systems Z_i and Z_j implies that also all the states of a composition are comparable: thus $(\forall z_i', z_i'', z_j', z_j'')(z_i'z_j' \Rightarrow z_i''z_j''$ or $z_i''z_j'' \Rightarrow z_i'\,z_j')$

This means that adiabatic isolation of the composition must be understood in a weak sense: the systems G_i and G_j are <u>together</u> adiabatically isolated with respect to their joint environment, but adiabatic walls between the two systems may be removed temporarily. In the third place: the metrization axiom

$$(\forall z_i', \ldots, z_i^n)\ (z_i' \ldots z_i^n \Rightarrow \text{Permutation of } z_i' \ldots z_i^n)$$

implies that we have to allow the use of additional systems, which describe cycle processes, within the isolation. E.g. consider two identical blocks of copper G_1 at a temperature of 100° C and G_2 at at temperature of 0° C. The axiom implies $z_1(100^\circ)\ z_2(0)^\circ \Rightarrow z_1(0^\circ)\ z_2(100^\circ)$. Such a process is however only possible with the help of for instance a carnot engine.

The latter property of the adiabatic isolation is more explicitly implied in the first metrization axiom

$$(\forall z_i', z_i'', z_j)\ (z_i'z_j \Rightarrow z_i''z_j \text{ iff } z_i' \Rightarrow z_j''),$$

where z_j may be interpreted as the initial and final state of a cyclic process $\langle z_j\ z_j \rangle$.

For certain systems $Z_i \in \mathcal{Z}$ the adiabatic isolation reduces to a symmetric relation: all the states of the system are reversibly adiabatically accessible. We interpret these systems as "mechanical systems". The entropy function $S_i(z_i)$ reduces to an arbitrary constant.

Δ

1.8 There exist for all $Z_i \in \mathcal{L}$ "energetic isolations", $(Iu)_i$ or $(u)_i$, being equivalence relations defined on $Z_i \times Z_i$, $Z_i \in \mathcal{L}$, and such that "extensivity or metrization axioms for internal energy" (axiom 3.3) hold.

1.9 There exist an energymeter Z_ν (definition 3.16) and all $Z_i \in \mathcal{L}$ are U-measurable (definition 3.19). with respect to this meter.

> This leads to the definition and existence proof of a set of extensive or additive internal energy functions $U_i\ (z_i)$ for the domains Z_i, $Z_i \in \mathcal{L}$.

Δ

We cannot introduce the thermodynamic concept "internal energy" in the usual way via the first law, if we intend to develop thermodynamics independent of mechanics. The concept of "work" is borrowed from mechanics and a definition is not possible without the help of an absolute force function and additive deformation coordinates. The solution chosen here is the introduction of the primitive concept "energetic isolation" as an equivalence relation which explains whether or not two states of a system "have the same internal energy". In traditional thermodynamics energetic isolation implies adiabatic isolation: thus traditional energetic isolation invokes an order in the set of equilibrium states of a system; to be precise: a partial preorder with the properties reflexivity and transitivity (the comparability of the simple preorder is thus lost.) The formal description of the energetic isolation as an equivalence relation assumes however symmetry, and an interpretation as energetic isolation in a traditional way is thus not suitable.

We shall interpret $z_i' \; z_i'' \in (Iu)_i$ or $z_i'(u)z_i''$ as 'the states z_i' and z_i'' are linked by processes possible under energetic isolation (in a traditional way) of the system G_i. '.

Δ

1.10 There exist for all $Z_i \in \mathcal{Z}$ "kth dimensional isolations" $(Ix_k)_i$, or $(x_k)_i$, being equivalence relations defined on $Z_i \times Z_i$, $Z_i \in \mathcal{Z}$, and such that "extensivity or metrization axioms for the kth deformation coordinate" hold.

1.11 There exists a meter $Z_{\xi k}$ for the kth deformation coordinate and all $Z_i \in \mathcal{Z}$ are X_k-measurable with respect to this meter.

> This leads to the definition and existence proof of a set of extensive or additive deformation coordinate functions $X_{ki}(z_i)$ for the domains Z_i, $Z_i \in \mathcal{Z}$.

Δ

The statement $z_i'z_i'' \in (Ix_k)_i$ or $z_i'(x_k)z_i''$ will be interpreted as: "the states z_i' and z_i'' are the outer states of a chain of possible processes under kth dimensional isolation, or under constant value of the kth deformation coordinate (e.g. constant volume)"

The concept of dimensional isolations is introduced as an interesting possibility to define extensive (additive) deformation coordinates in a similar way as we defined extensive (additive) entropy and internal energy variables.

Whether this approach has more than a formal significance may be doubted. If preferred one may replace this set of axioms by the straightforward statement of the existence of a set of additive (extensive) deformation coordinates X_{ki}:

Δ or

1.10* There exist extensive or additive deformation coordinate functions $X_{ki}(z_i)$ defined on $Z_i \in \mathcal{Z}$

Δ

Looking in retrospect to the above remarks it will be clear that they cannot be considered as rules of interpretation. In a certain sense the question of interpretation is begged. We have associated the primitive terms of the formal theory with terms of the physical theory but not answered such questions as: how to decide whether a physical object can be considered as a closed system; how to decide whether a relationship between physical systems is a thermal connection, a force connection, or that a physical system is adiabatically, energetically, or kth dimensionally isolated. These questions lie however outside the scope of this work and I will not try to give a provisional answer, aware of the considerable difficulties into which we are led.

Extremal principles, absolute temperature and absolute forces

---

Δ

1.12 The systems $Z_i \in \mathcal{Z}_\theta$ have a phase space $[\,U_i, \ldots\ldots X_{ki} \ldots]$, $k \in N_i \subset N$, where N is the set of integers which indicate the different kinds of deformation coordinates (axiom 4.3)

Δ

The axiom is usually either tacitly assumed or an implication of other assumptions. For instance, it follows from the assumption that systems $Z_i \in \mathcal{Z}_\theta$ have a phase space $[\theta, \ldots X_{ki} \ldots]$ and that U_i is a monotonic increasing function of the empirical temperature θ.

The importance of this axiom lies in the selection of a preferred set of variables for systems $Z_i \in \mathcal{Z}_\theta$, which are interpreted as systems without adiabatic partitions. The set of variables $[\,U_i, \ldots X_{ki} \ldots\,]$, which are all extensive, plays a crucial role in the further development of the theory and difficulties which arise through the choice of other sets of variables, e.g. $[\theta, \ldots X_{ki} \ldots]$, caused bydiscontinuities or kinks in the $S(\theta \ldots X_{ki} \ldots)$ or $\mathcal{U}(\theta, \ldots X_{ki} \ldots)$ functions can be avoided.

Δ

1.13 For systems $Z_i \in \mathcal{Z}_\theta$ the maximum entropy principle (axiom 5.1) holds.

This leads under certain "local continuity assumptions" (theorem 5.11) to the definition of an absolute negative reciprocal temperature function $N(z_i) \equiv - {}^1/T(z_i)$ (definition 5.13)

1.14 The systems $Z_i \in \mathcal{Z}_\theta$ are measurable with a set of absolute thermometers, covering together the temperature domain $- \infty < N < + \infty$, or a subdomain (axiom 5.15); an absolute thermometer being a system of $\mathcal{Z}_\theta$ for which certain "continuity assumptions" hold.

Δ

The maximum entropy principle, introduced in thermodynamics by F.W. Gibbs, is by many authors recognised as a powerful axiomatic starting point.
Its precise formulation has given considerable difficulties. Tisza was presumably the first to give an unambiguous formulation. Our formulation is adapted to the theory, as presented here, but closely related to Tisza's.

Δ

1.15 For systems $Z_i \in \mathcal{Z}_{\phi_k}$ the "mechanical" minimum energy principle (axiom 5.2) holds.

This leads under certain "local continuity (theorem 5.12) assumptions"; to the definition of an absolute force function $F_k(z_i)$ (definition 5.14)

1.16 The systems $Z_i \in \mathcal{Z}_{\phi_k}$ are measurable with a set of absolute dynamometers for force k, covering together the force domain $-\infty < F_k < +\infty$, or a subdomain; an absolute dynamometer force k being a system of $\mathcal{Z}_{\phi_k}$ for which certain "global continuity assumptions" hold (axiom 5.16)

Δ

In a development of thermodynamics, which does not presuppose mechanics, a minimum internal energy principle is no less important than the maximum entropy principle. It has in its formulation striking similarities with the latter principle. It leads to the introduction of an absolute force concept, which can be interpreted as the force, usually borrowed from mechanics. Its definition is of a static character, it does not presuppose newtonian dynamics or similar theories. Together with the maximum entropy principle it yields the Gibbs fundamental equation for closed systems:

$$dS_i = (1/T)\,d\mathcal{U}_i - \sum_k (1/T)\,F_k\,d\,X_{ki}.$$

The mechanical minimum energy principle must be distinguished from the Gibbsian minimum energy principle, which can be proved equivalent to the maximum entropy principle within the framework of the theory.

Δ

1.17 For all $Z_i \in \mathcal{Z}$, $S_i(z_i)$ has a lower bound

(axiom 6.4)

This leads in the case of systems for which "global continuity assumptions" obtain, to the existence of bounds for the energy and deformation coordinates (theorem 6.5) and to certain properties of the systems concerned at the boundaries of the domain of definition in the phase space $[U_i, \ldots X_{ki} \ldots]$ (theorem 6.6).

1.18 For all $Z_i \in \mathcal{Z}$: $\lim\limits_{N \to -\infty} S_i (U_i \ldots X_{ki} \ldots)$ is independent of U_i and X_{ki} (all $k \in N_i$) or

$$\lim_{\substack{T(z_i) \to 0 \\ N(z_i) < 0}} \left[\frac{\partial S_i(z_i)}{\partial U_i} \right]_T = 0$$

$$\lim_{\substack{T(z_i) \to 0 \\ N(z_i) > 0}} \left[\frac{\partial S_i(z_i)}{\partial X_i} \right]_T = 0$$

Δ

These axioms comprise all that still has to be stated to cover the traditional content of the third law. Different aspects of the third law appear already to be consequences of the foregoing axioms.

CHAPTER 2 : ENTROPY

In this and the following two chapters I will develop mainly the formal theory. This will be done in a series of axioms, definitions and theorems, which again are considered as self contained. It appears however profitable to embed this core in a text, that explains the main lines of thought and stresses the more important results. This text belongs also to the formal theory. We do not need special signs to separate axioms, definitions and theorems from the embedding remarks, except that we mark the end of a proof with the sign $\square$. In the few places where we leave the formal theory, for instance to explain the physical interpretation of a certain statement or symbol, I will again use horizontal lines and the sign Δ.

Δ

The aim of this chapter is to construct a set of extensive entropy functions $S_i(z_i)$ for the systems $Z_i \in \mathcal{Z}$.

2.1 Def. A set of extensive entropy functions S_i for the systems $Z_i \in \mathcal{Z}$ is a set of real valued functions $S_i(z_i)$, $z_i \in Z_i$, such that

(i) $S_i(z_i') \leq S_i(z_i'')$ iff $z_i' \Rightarrow z_i''$

(ii) if $Z \equiv Z_i \times Z_j \times$ ___ then

$$S(z) \equiv S(z_i\ z_j\ \ldots.) = S_i(z_i) + S_j(z_j) + \text{---}$$

For the construction of such a set of functions, conditions of different kinds must be fulfilled.

The first kind consists of relationships between the adiabatic isolation relations $\Rightarrow_{ijk\,\ldots}$ of the composition

$Z_{ijk} \equiv Z_i \times Z_j \times Z_k \times$ ___ and the adiabatic isolation relations $\Rightarrow i$, $\Rightarrow_j$, $\Rightarrow_k$, ___ of the individual component systems Z_i, Z_j, Z_k, ----. We call these relationships "metrization axioms for entropy".

The second kind contains only an existence statement with respect to a system with special properties: an entropymeter Z_σ
A procedure for the construction of a function $S_\sigma(z_\sigma)$ and of functions $S_i(z_i)$ is given and it is proved that this set of functions has the properties of extensive entropy functions. Starting with a detailed presentation of this part of the theory we define first two useful relations in terms of the relation $\Rightarrow$.

2.2 Def. $z' \rightarrow z''$ iff $z' \Rightarrow z''$ and not $z'' \Rightarrow z'$

2.3 Def. $z' \leftrightarrow z''$ iff $z' \Rightarrow z''$ and $z'' \Rightarrow z'$

Δ

The interpretation of these relations will be clear:
$z' \rightarrow z''$ has to be interpreted as : z'' is irreversible adiabatically accessible from z', and $z' \leftrightarrow z''$ means z' and z'' are connected by a reversible adiabatic process. The term adiabatic process has to be interpreted as explained in Chapter 1.

Δ

The following four theorems give properties of these relations which we will use in the further development of the theory:

2.4 Th. The relation $\rightarrow$ is transitive and asymmetric
(i.e. if $z' \rightarrow z''$ then not $z'' \rightarrow z'$)

Proof:

transitivity:
$z' \rightarrow z''$ and $z'' \rightarrow z'''$, iff (2) $z' \Rightarrow z''$ and $z'' \Rightarrow z'''$ and not $z'' \Rightarrow z'$ and not $z''' \Rightarrow z''$, iff $z' \Rightarrow z'''$ and not $z''' \Rightarrow z'$ (because

not $z'' \Rightarrow z'$ and not $z''' \Rightarrow z''$ implies not $z''' \Rightarrow z'$.
If not, then ($\exists z'z''z''$) not $z'' \Rightarrow z'$ and not $z''' \Rightarrow z''$ and $z''' \Rightarrow z'$, thus, as $\Rightarrow$ is comparable, not $z'' \Rightarrow z'$ and $z'' \Rightarrow z'''$ and $z''' \Rightarrow z'$, thus, as $\Rightarrow$ is transitive, not $z'' \Rightarrow z'$ and $z'' \Rightarrow z'$, a contradiction).

Asymmetry follows immediately from the definition $\square$

2.5 Th. The relation $\leftrightarrow$ is transitive, reflexive and symmetric
This follows immediately from the definition.
2.6 Def: The equivalence classes of $\leftrightarrow$ are called the "classes of states of equal entropy".
2.7 Th. if $z' \rightarrow z''$ and $z' \leftrightarrow z'''$ then $z''' \rightarrow z''$.
Proof:
if not then ($\exists z'z''z'''$) $z' \rightarrow z''$ and $z' \leftrightarrow z'''$ and not $z''' \rightarrow z''$;
not $z''' \rightarrow z''$ implies (2) not $z''' \Rightarrow z''$ or $z'' \Rightarrow z'''$
$z' \rightarrow z''$ and $z' \leftrightarrow z'''$ implies (2,3) $z' \Rightarrow z''$ and $z''' \Rightarrow z'$ and thus $z''' \Rightarrow z''$.
$z'' \Rightarrow z'''$ and $z' \leftrightarrow z'''$ imply $z'' \Rightarrow z'''$ and $z''' \Rightarrow z'$ and thus $z'' \Rightarrow z'$ contradictory with $z' \rightarrow z''$. $\square$

2.8 Th. if $z' \Rightarrow z''$ and $z' \leftrightarrow z'''$ then $z''' \Rightarrow z''$.

We next formulate the <u>metrization axioms for entropy</u>:

2.9 Ax. For systems $Z_i, Z_j \in \mathcal{Z}$ the following holds:
(i) $(\forall\ z_i', z_i'', z_j)$ $(z_i' \Rightarrow z_i''$ iff $z_i'z_j \Rightarrow z_i''z_j)$
(ii) $(\forall z_i^1, \dots z_i^n)$ $(z_i^1 \dots z_i^n \Rightarrow$ Permutation $z_i^1 \dots z_i^n)$
(iii) $(\forall z_i', z_i'')$ (if $(z_i' \dots z_i')^n \Rightarrow (z_i'' \dots z_i'')^n$ then

$z_i' \Rightarrow z_i''$) where $(z_i' \dots z_i')^n$ means a state of a composition of n identical systems Z_i consisting of n identical states z_i'. The brackets and the index n are omitted many times, because there is never any doubt about the number of states z_i' and z_i'' involved.

It is not difficult to show that the sign $\Rightarrow$ in these axioms may be replaced by $\leftrightarrow$ and in the case of (i) and (iii) also by $\rightarrow$. This is the content of theorem 11 and theorem 13. For the proofs we need some additional theorems.

2.10 Th. $(z_i' \dots z_i')^n \Rightarrow (z_i'' \dots z_i'')^n$ iff $z_i' \Rightarrow z_i''$

Proof:

if $z_i' \Rightarrow z_i''$ then $z_i'z_i' \Rightarrow z_i''z_i'$ and $z_i'z_i'' \Rightarrow z_i''z_i''$ thus $z_i'z_i' \Rightarrow z_i''z_i''$ etc. Axiom 9.iii gives the other half of the proof. □

2.11 Th. For arbitrary systems Z_i, Z_j

(i) $(\forall z_i', z_i'', z_j)$ $(z_i' \leftrightarrow z_i''$ iff $z_i'z_j \leftrightarrow z_i''z_j)$

(ii) $(\forall z_i^1, \dots z_i^n)$ $(z_i^1 \dots z_i^n \leftrightarrow$ permutation $z_i^1 \dots z_i^n)$

(iii) $(\forall z_i', z_i'')$ $((z_i' \dots z_i')^n \leftrightarrow (z_i'' \dots z_i'')^n$ iff $z_i' \leftrightarrow z_i'')$

Proof:

immediate consequences of the definition of $\leftrightarrow$ and the axiom 9 and theorem 10 □

2.12

Th. if $z_1' \dots z_i' \dots z_n' \leftrightarrow z_1'' \dots z_i'' \dots z_n''$ and $z_i' \leftrightarrow z_i'''$ then $z_1' \dots z_i''' \dots z_n' \leftrightarrow z_1'' \dots z_i'' \dots z_n''$.

Proof:

$z_i' \leftrightarrow z_i'''$ implies (11.i,11.ii) $z_1' \dots z_i' \dots z_n' \leftrightarrow z_1' \dots z_i''' \dots z_n'$ This together with $z_i' \dots z_i' \dots z_n' \leftrightarrow z_1'' \dots z_i'' \dots z_n''$ and the transitivity of $\leftrightarrow$ yields the theorem. □

2.13 Th. For arbitrary systems $Z_i, Z_j, \in \mathcal{Z}$

(i) $(\forall z_i', z_i'', z_j)$ $(z_i' \rightarrow z_i''$ iff $z_i'z_j \rightarrow z_i''z_j)$

(ii) $(\forall z_i', z_i'')$ $\left[(z_i' \dots z_i')^n \rightarrow (z_i'' \dots z_i'')^n \text{ iff } z_i' \rightarrow z_i''\right]$

(an immediate consequence of the definition of $\rightarrow$ and axiom 9 and theorem 10)

The next two theorems will be needed in the second part of this chapter:

2.14 Th: If $z_1' \rightarrow z_1''$ and $z_2' \rightarrow z_2''$ and ....

$z_n' \rightarrow z_n''$ then $z_1' \dots z_n' \rightarrow z_1'' \dots z_n''$.

Proof:

(13.i) $z_1' \rightarrow z_1''$ iff $z_1' z_2' \rightarrow z_1'' z_2'$

(13.i) $z_2' \rightarrow z_2''$ iff $z_2' z_1'' \rightarrow z_2'' z_1''$

Thus $z_1' \rightarrow z_1''$ and $z_2' \rightarrow z_2''$ implies (11.ii, and transitivity of $\rightarrow$), $z_1' z_2' \rightarrow z_1'' z_2''$.

2.15 Th. if $z' \rightarrow z''$ and $z' z'' \Leftrightarrow z\ z$, then $z' \rightarrow z \rightarrow z''$

Proof:

(13.i) $z' \rightarrow z''$ implies $z' z' \longrightarrow z'' z'$ and $z' z'' \rightarrow z'' z''$, thus (4,7) $z' z' \rightarrow z' z'' \rightarrow z'' z''$, thus (7) $z' z' \rightarrow z\ z \rightarrow z'' z''$, thus (13.iii) $z' \rightarrow z \rightarrow z''$.

The entropy meter and its calibration

We define, in a certain sense simultaneously, a special system Z_σ, called "entropy meter and a function $S_\sigma (z_\sigma)$." We shall start with the latter, because it will clarify the properties of the meter.

2.16 Def. A real valued function $S(z_\sigma)$ for $z_\sigma \epsilon\ Z_\sigma$ is constructed as follows:

Choose two states z'_σ, $z''_\sigma \in Z_\sigma$ such that $z'_\sigma \rightarrow z''_\sigma$
$S(z'_\sigma) \equiv 0$, $S(z''_\sigma) \equiv 1$, we write $z'_\sigma = z_\sigma(0)_s$, $z''_\sigma = z_\sigma(1)_s$
$S(z_\sigma) = \frac{1}{2}$ iff $z_\sigma(0)_s\, z_\sigma(1)_s \leftrightarrow z_\sigma z_\sigma$ etc.
$S(z_\sigma) = n$(integer) iff $z_\sigma z_\sigma(n-2)_s \leftrightarrow z_\sigma(n-1)_s\; z_\sigma(n-1)_s$
$S(z_\sigma) = -1$ iff $z_\sigma z_\sigma(1)_s \leftrightarrow z_\sigma(0)\, z_\sigma(0)$ etc.
$S(z_\sigma) = r$ (real) iff $\sup \{ d : z_\sigma(d) \Rightarrow z_\sigma \} = r$ or
$\inf \{ d : z_\sigma \Rightarrow z_\sigma(d) \} = r$

(In case that $S(z_\sigma) = n$ and there does not exist a z^*_σ such that $S(z^*_\sigma) = n+1$ we define : $S(z_\sigma) = n + \frac{1}{2}$ iff
$z_\sigma z_\sigma(n - \frac{1}{2})_s \leftrightarrow z(n)\, z(n)$.)

The above procedure can justly be called a <u>calibration</u> of the meter Z_δ.

2.17 Def. An entropymeter is a system Z_σ such that

(i) $(\mathcal{S}_\sigma \rightarrow\ ) \simeq (\mathcal{R}', <)$, $\mathcal{R}'$ being an interval of reals and $\mathcal{S}_\sigma$ being the family of equivalence classes

(ii) "calibration property":

$$(\forall z_\sigma)(\forall \mathcal{N}_\delta(z_\sigma))\ (\exists z_\sigma(d')_s,\ z_\sigma(d'')_s)$$
$$\left[z_\sigma(d')_s,\ z_\sigma(d'')_s \in \mathcal{N}_\delta(z_\sigma) \text{ and } z_\sigma(d')_s \rightarrow z_\sigma \rightarrow z_\sigma(d'')_s \right]$$

where a neighborhood $\mathcal{N}_\delta(z_\sigma)$ is defined as a set $\{z^\dagger_\sigma : z'_\sigma \rightarrow z^\dagger_\sigma \rightarrow z''_\sigma\}$ containing z_σ.

(iii) $(\forall z'_\sigma, z''_\sigma)(\exists z_\sigma)\ (z'_\sigma z''_\sigma \leftrightarrow z_\sigma z_\sigma)$

We explain the formalism of (i) as follows: there exists a similarity mapping from the family of classes of states of equal entropy, ordered through the relation → onto an interval of real numbers $\mathcal{R}'$ in their natural order $<$. (ii) can be worded in this way: every interval $\{z^*_\sigma : z_\sigma' \rightarrow z^*_\sigma \rightarrow z_\sigma''\}$ containing z_σ , contains also states with dyadic S_σ values $z_\sigma(d')_s$ and $z_\sigma(d'')_s$ such that $z_\sigma(d')_s \rightarrow z_\sigma \rightarrow z_\sigma(d'')_s$. Because it is not immediately clear that (ii) and (iii) are mutually independent, we prove this by showing that a system can be defined for which ax.9 and properties (i) and (ii) obtain and for which (iii) is falsified and a system for which ax.9 and properties (i) and (iii) obtain and (ii) is falsified.

Suppose a system Z such that 17 (i) and (ii) obtains and such that the composition $Z \times Z$ satisfies:

if $r_1 + r_2 < r_3 + r_4$ then $z(r_1)\ z(r_2) \rightarrow z(r_3)\ z(r_4)$;

if $r_1 + r_2 = r_3 + r_4 = d$ then $z(r_1)\ z(r_2) \leftrightarrow z(r_3)\ z(r_4)$;

if $r_1 + r_2 = r_3 + r_4 \neq d$ then

if $(r_1 - r_2)^2 < (r_3 - r_4)^2$ then $z(r_1)\ z(r_2) \rightarrow z(r_3)\ z(r_4)$

if $(r_1 - r_2)^2 = (r_3 - r_4)^2$ then $z(r_1)\ z(r_2) \leftrightarrow z(r_3)\ z(r_4)$

We can check that ax. 9 is satisfied in this case. However, if $r_1 + r_2 \neq d$ and $r_1 \neq r_2$ then there exists no state $z(r)$ such that $z(r_1)\ z(r_2) \leftrightarrow z(r)\ z(r)$

Remark: in this example however,

$(\forall z(d'), z(d''))\ (\exists z(d))\ (z(d')\ z(d'') \leftrightarrow z(d)\ z(d))$

It can be proved that this is a consequence of property (ii). A strengthening of property (ii) in the following sense makes (iii) a theorem:

(ii-a) $(\forall z_\sigma', z_\sigma'')\ (\forall \mathcal{N}_s(z_\sigma'\ z_\sigma''))\ (\exists z_\sigma(d_1), z_\sigma(d_2), z_\sigma(d_3), z_\sigma(d_4))$
$\{z_\sigma(d_1)\ z_\sigma(d_2),\ z_\sigma(d_3)\ z_\sigma(d_4) \in \mathcal{N}_s(z_\sigma'\ z_\sigma'')$ &
$z_\sigma(d_1)\ z_\sigma(d_2) \rightarrow z_\sigma'\ z_\sigma'' \rightarrow z_\sigma(d_3)\ z_\sigma(d_4)\}$.

An example of a system for which 17 (i) and (iii) obtains but not 17 (ii) is the following.

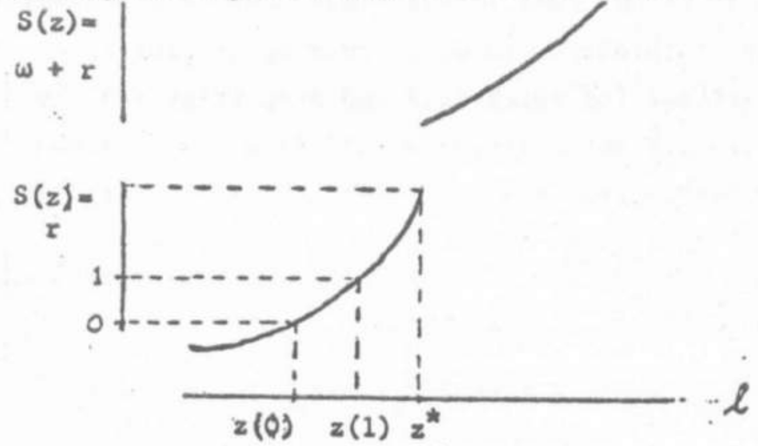


figure 2.1

Property (i) allows us to map the system $\mathcal{S}$, → onto a line element ℓ, while preserving the order. Suppose the maps of $z(0)$ and $z(1)$ lie on the line ℓ as indicated in the figure. It is quite possible that in the construction of the function $S(z)$ the procedure fails for all states z such that $z^* \rightarrow z$. We are forced in this case to the conclusion $S(z) = S(z^*)$ (= say r) for all $z : z^* \rightarrow z$, thus $S(z)$ is not an entropy function. This situation is incompatible with property (ii), and the maintenance of property (iii) does not lead to inconsistencies. Suppose for instance that in the domain $\{z : z^* \rightarrow z\}$, choosing two states $z(0)'$, $z(1)'$ an $S'(z)$ function can be constructed which is real for all z in this domain, and that we unify the two real valued functions for the respective domains $\{z : z \Rightarrow z^*\}$ and $\{z : z^* \rightarrow z\}$ to one function $S(z)$ which takes in the first domain the values $S(z) = r$ and in the second the values $S(z) = \omega + r$ (thus not real). It is possible to associate with a pair $z(r_1)z(\omega + r_2)$ a unique value $\omega + r_3$ such that $z(r_1)z(\omega + r_2) \leftrightarrow z(\omega + r_3)z(\omega + r_3)$ without inconsistencies with axiom 9 and properties 17 (i) and (iii)

On the basis of the metrisation axioms for entropy and the definition of the entropy meter Z_σ and the function $S_\sigma(z_\sigma)$ we can prove the following two theorems, the first stating that the constructed function is an entropy function with certain continuity characteristics, the second that the function for compositions of identical systems Z_σ behaves as an <u>extensive</u> entropy function.

2.18 Th. For an entropy meter Z_σ:

(a) $\{r : S(z_\sigma) = r\} = \mathcal{R}'$ where $\mathcal{R}'$ is an interval of reals

(b) $(\forall z'_\sigma, z''_\sigma)\ (S(z'_\sigma) < S(z''_\sigma)$ iff $z' \rightarrow z'')$

Proof:

(a) Let $\mathcal{R}''$ be the smallest interval of reals which contains $\{r : S(z_\sigma) = r\}$. (17 (iii)) guarantees that for all dyadics $d = p/_2g$, p,g being integers, if $d \in \mathcal{R}''$ then there exists a state z_σ such that $S(z_\sigma) = d$.

The construction of the function $S(z_\sigma)$ guarantees that for the states with dyadic S values the natural order according to these values is also an entropic order. This is a consequence of Theorem 15. Thus, if $d' < d''$ then $z_\sigma(d')_s \rightarrow z_\sigma(d'')_s$.

The calibration property 17.ii guarantees that every z_σ divides the entropically ordered set $\{z_\sigma : S(z_\sigma) = d\}$ in a lower and an upper cut, thus for all z_σ $\sup\{d : z_\sigma(d)_s \rightarrow z_\sigma\} = \inf\{d : z_\sigma \rightarrow z_\sigma(d)_s\} = r$. Finally, if $r \in \mathcal{R}''$ then $(\exists z_\sigma)\ (S(z_\sigma) = r)$. If not, thus if an entropy value r is "missing" then we can divide the set $(Z_\sigma, \rightarrow)$ into two succeeding subsets $Z_{\sigma 1} \equiv \{z_\sigma : z_\sigma \Rightarrow z_\sigma(d)_s,\ d < r\}$ and $Z_{\sigma 2} \equiv \{z_\sigma : z_\sigma(d)_s \Rightarrow z_\sigma,\ r < d\}$

$Z_{\sigma 1}$ does not contain a last state, $Z_{\sigma 2}$ does not contain a first state, and $Z_{\sigma 1} \cup Z_{\sigma 2} = Z_\sigma$. With respect to the mapping $(Z_\sigma, \rightarrow) \leftrightarrow (\mathcal{R}, <)$ this implies that also $(\mathcal{R}, <)$ can be divided in two succeeding ordered sets $(\mathcal{R}_1, <)$ and $(\mathcal{R}_2, <)$ such that

$\mathcal{R}_1$ has no last element, $\mathcal{R}_2$ no first element, and $\mathcal{R}_1 \cup \mathcal{R}_2 = \mathcal{R}$.
This contradicts the supposition (17.1)

(b) We prove first : if $z'_\sigma \to z''_\sigma$ then $S(z'_\sigma) < S(z''_\sigma)$
If not, then $z'_\sigma \to z''_\sigma$ and $S(z'_\sigma) \geqslant S(z''_\sigma)$;

$z'_\sigma \to z''_\sigma$ and $S(z'_\sigma) = S(z''_\sigma)$ implies that all the states z_σ in the interval $\{z_\sigma : z'_\sigma \to z_\sigma \to z''_\sigma\}$ have the same value $S(z_\sigma) = S(z'_\sigma) = S(z''_\sigma)$ which contradicts the calibration property.
$z'_\sigma \to z''_\sigma$ and $S(z'_\sigma) > S(z''_\sigma)$ implies $(\exists d)(S(z''_\sigma) < d < S(z'_\sigma))$ and thus, because of the definition of $S(z_\sigma)$, $z''_\sigma \to z_\sigma(d)$ and $z_\sigma(d) \to z'_\sigma$, thus $z''_\sigma \to z'_\sigma$, contradictory to the supposition.
The same reasoning leads to : if $S(z'_\sigma) < S(z''_\sigma)$ then $z'_\sigma \to z''_\sigma$ □

2.19 Th. For the compositions of identical entropymeters Z_σ :
$z_\sigma(r_1)\, z_\sigma(r_2) \dots z_\sigma(r_m) \Rightarrow z_\sigma(r'_1)\, z_\sigma(r'_2) \dots z_\sigma(r'_m)$ iff
$\sum_1^m r_k \leqslant \sum_1^m r'_k$

Proof:
We prove the theorem first for dyadic rationals in the interval (0,1), then for all dyadics and finally for all the reals.
We omit subscripts, writing $z(r)$ in place of $z_\sigma(r)_s$.

(i) Dyadics in (0,1) can be written in a form $p/2^n$,

$p \in \{(0,1, \dots 2^n\}$. If $n = 0$ then the lemma
$z(d_1) \dots z(d_m) \Rightarrow z(d'_1) \dots z(d'_m)$ iff $\sum_1^m d_m \leqslant \sum_1^m \leqslant d'_m$

is clearly true.
We prove that if the lemma is true for $n - 1$ then it is true for n, or if it is true for dyadics of the form $p/2^{n-1}$, $p \in \{0,1, \dots 2^{n-1}\}$ then it is true for dyadics of the form $q/2^n$, $q \in \{0,1, \dots 2^n\}$.

If q is even then we write q_i $\quad {}^{q_i}/_{2^n} = {}^{p_i}/_{2^{n-1}}$

If q is odd then we write q_j $\quad {}^{q_j}/_{2^n} = (2p_j + 1)/_{2^n}$

Now (10):

$$z\left[{}^{q_1}/_{2^n}\right] \ \ldots\ z\left[{}^{q_m}/_{2^n}\right] \Rightarrow z\left[{}^{q_1'}/_{2^n}\right] \ \ldots\ z\left[{}^{q_m'}/_{2^n}\right] \quad \text{iff}$$

$$z\left[{}^{q_1}/_{2^n}\right]\ z\left[{}^{q_1}/_{2^n}\right] \ \ldots\ \cancel{z\left[{}^{q_m}/_{2^n}\right] \Rightarrow}\ z\left[{}^{q_m}/_{2^n}\right]\ z\left[{}^{q_m}/_{2^n}\right] \Rightarrow$$
$$z\left[{}^{q_1'}/_{2^n}\right]\ z\left[{}^{q_1'}/_{2^n}\right] \ \ldots\ z\left[{}^{q_m'}/_{2^n}\right]\ z\left[{}^{q_m'}/_{2^n}\right]$$

$$\text{iff} \quad \left[\ldots\ldots\ z\left[{}^{q_i}/_{2^n}\right]\ z\left[{}^{q_i}/_{2^n}\right] \ldots\ z\left[{}^{q_j}/_{2^n}\right]\ z\left[{}^{q_j}/_{2^n}\right] \ldots \right] \Rightarrow$$
$$\left[\ldots\ldots\ z\left[{}^{q_i'}/_{2^n}\right]\ z\left[{}^{q_i'}/_{2^n}\right] \ldots\ z\left[{}^{q_j'}/_{2^n}\right]\ z\left[{}^{q_j'}/_{2^n}\right] \ldots \right] \text{iff}$$

(using the definition of dyadics)

$$\left[\ldots\ z\left[{}^{p_i}/_{2^{n-1}}\right]\ z\left[{}^{p_i}/_{2^{n-1}}\right] \ldots\ z\left[{}^{q_j+1}/_{2^n}\right]\ z\left[{}^{q_j-1}/_{2^n}\right] \ldots \right] \Rightarrow$$
$$\left[\ldots\ z\ {}^{p_i'}/_{2^{n-1}}\right]\ z\left[{}^{p_i'}/_{2^{n-1}}\right] \ldots\ z\left[{}^{q_j'+1}/_{2^n}\right]\ z\left[{}^{q_j'-1}/_{2^n}\right] \ldots \right] \quad \text{iff}$$
$$\left[\ldots\ z\left[{}^{p_i}/_{2^{n-1}}\right]\ z\left[{}^{p_i}/_{2^{n-1}}\right] \ldots\ z\left[{}^{p_j+1}/_{2^{n-1}}\right] z\left[{}^{p_j}/_{2^{n-1}}\right] \ldots \right] \Rightarrow$$
$$\left[\ldots\ z\left[{}^{p_i'}/_{2^{n-1}}\right]\ z\left[{}^{p_i'}/_{2^{n-1}}\right] \ldots\ z\left[{}^{p_j'+1}/_{2^{n-1}}\right] z\left[{}^{p_j'}/_{2^{n-1}}\right] \ldots \right]$$

iff (supposing the lemma obtains for n-1)

$$\Sigma(p_i + p_i) + \Sigma(p_j + p_j + 1) \leq \Sigma(p_i' + p_i') + \Sigma(p_j' + p_j' + 1)$$

iff $\underset{\text{all even q's}}{\Sigma q_i} + \underset{\text{all odd q's}}{\Sigma q_j} \leq \underset{\text{all even q's}}{\Sigma q_i'} + \underset{\text{all odd q's}}{\Sigma q_j'}$

iff $\sum_1^m d_k \leq \sum_1^m d_k'$

(ii) The extension of the theorem to all dyadics starts with the proof that it is true for $n = 0$, $p \in \{0, 1, 2\}$

$z(p_1) \;\ldots\ldots\; z(2) \;\ldots\; z(p_m) \Rightarrow z(p_1') \;\ldots\; z(p_m')$ iff

$z(p_1) \;\ldots\ldots\; z(2) \;\ldots\; z(p_m)\; z(0) \Rightarrow z(p_1') \;\ldots\; z(p_m')\; z(0)$ iff

$z(p_1) \;\ldots\ldots\; z(1) \;\ldots\; z(p_m)\; z(1) \Rightarrow z(p_1') \;\ldots\; z(p_m')\; z(0)$

and this reduces this case to the case $n = 0$, $p \in \{0, 1\}$

In the same way we extend it to $p \in \{0, 1, 2, 3\}$ etc.

The rest of (i) was independent of the value of p.

(iii) The extension to the reals starts with the deduction of:

if $z(r_1) \ldots z(r_m) \Rightarrow z(r_1') \ldots z(r_m')$ then $\sum_1^m r_k \leq \sum_1^m r_k'$

If not then $z(r_1) \ldots z(r_m) \Rightarrow z(r_1') \ldots z(r_m')$ and

$\sum_1^m r_k' < \sum_1^m r_k$ Then there exists a dyadic d such that $\Sigma r_k' < d < \Sigma r_k$,

and sets of dyadics $\{d_k, \ldots\}$ and $\{d_k', \ldots\}$ such that

$$d = \sum_1^m d_k = \sum_1^m d_k' \text{ and } r_k' < d_k' \text{ and } d_k < r_k.$$

This implies $z(r_k') \rightarrow z(d_k')$ and $z(d_k) \rightarrow z(r_k)$ and

Consequently (14) $z(r_1') \ldots z(r_m') \rightarrow z(d_1') \ldots z(d_m')$ and

$z(d_1) \ldots . z(d_m) \quad \rightarrow \quad z(r_1) \ldots . z(r_m)$.

Because $\Sigma d_k = \Sigma d_k' \quad z(d_i') \ldots . z(d_m') \Leftrightarrow z(d_1) \ldots . z(d_m)$

and thus (7) $z(r_1') \ldots z(r_m') \rightarrow z(r_1) \ldots . z(z_m)$

contradictory to the supposition.

Finally we prove:

If $\sum_1^m r_k \leq \sum_1^m r_k'$ then $z(r_1) \ldots z(r_m) \Rightarrow z(r_1') \ldots . z(r_m')$

If $\sum_1^m r_k < \sum_1^m r_k'$ then again there exist sets of dyadics

$\{d_k \ldots\}$ and $\{d_k'\}$ such that $\Sigma r_k < \Sigma d_k = \Sigma d_k' < \Sigma r_k'$ and $r_k < d_k$ and $d_k' < r_k'$. It follows that $z(r_1) \ldots z(r_m) \rightarrow$

$z(r_1) \ldots z(r_m')$. If $\sum_1^m r_k = \sum_1^m r_k$ then $z(r_1) \ldots z(r_m) \leftrightarrow z(r_1') \ldots z(r_m')$ will be proved next:

Consider the state $z(r_1) \ldots . z(r_m) \; z(0) \ldots . z(0)$, in total 2^ℓ elements z. By ℓ times application of $(\exists z)\,(z' z'' \Leftrightarrow z\, z)$ we find:

$\exists z(r) \quad [z(r_1) \ldots z(r_m)\, z(0) \ldots z(0) \Leftrightarrow (z(r) \ldots z(r)^{2\ell}]$

But the first part of (19 iii) gives then $\sum_1^m r_k = 2^\ell . r$

Similarly we find

$\exists z(r') \quad [z(r_1') \ldots z(r_m')\, z(0) \ldots z(0) \Leftrightarrow (z(r) \ldots z(r))^{2\ell}]$

$\sum_1^m r_k' = 2^\ell r'$.

But we suppose that $\Sigma r_k = \Sigma r_k'$, thus $r = r'$, thus

$z(r_1) \ldots z(r_m)\ z(0) \ldots z(0) \longleftrightarrow z(r_1') \ldots z(r_m')\ z(0) \ldots z(0)$

thus $z(r_1) \ldots z(r_m) \longleftrightarrow z(r_1') \ldots z(r_m')$

The following step is the definition of a function $S_i(z_i)$ for an arbitrary system $Z_i \in \mathcal{Z}$ with the help of an entropy meter Z_σ. The construction can be called a "<u>measuring procedure</u>".

2.20 Def. A function $S_i(z_i)$ is constructed as follows:

Choose an entropy meter Z_σ and a state $z_i' \in Z_i$

$$S_i(z_i') \equiv 0 \quad \text{we write } z_i' = z_i(0)_s$$

$$S_i(z_i) = r_i \quad \text{iff} \quad z_i\ z_\sigma(0) \longleftrightarrow z_i(0)_s\ \ z_\sigma(r)_s$$

(or more generally)

$$S_i(z_i) = r \quad \text{iff} \quad (z_i\ z_\sigma(r')_s \ldots z_\sigma(r')_s)^{n+1} \longleftrightarrow (z_i(0)_s z_\sigma(r'') \ldots z_\sigma(r''))^{n+1}$$

and $n(r'' - r') = r$)

There is still no guarantee that for all states of all systems real values $S_i(z_i)$ can be defined, with respect to a chosen entropy meter Z_σ if for a system Z_i this condition is fulfilled we call the system S-measurable with respect to the meter Z_σ.

2.21 Def. The system Z_i is S-measurable with the entropy meter Z if for every $z_i \in Z_i$ $S(z_i) = r$ (r being a real number).

The following axiom guarantees that all systems $Z_i \in \mathcal{Z}$ are S-measurable.

2.22 Ax. All systems $Z_i \in \mathcal{Z}$ are S-measurable with respect to the entropy meter Z_σ

The existence proof of an extensive entropy function defined on all states $z_i \in Z_i$, $Z_i \in \mathcal{Z}$ is straightforward.

2.23 Th. For a set of systems $Z_1 \ldots\ldots Z_m \in \mathcal{Z}$ obtains for all the possible compositions:

$$z_1 \ldots\ldots\ldots\ldots z_m \Rightarrow z'_1 \ldots\ldots\ldots\ldots z'_m \quad \text{iff} \quad \sum_1^m S_i(z_i) \leq \sum_1^m S_i(z'_i)$$

Proof

$$(9.i) \quad z_1 \ldots z_m \Rightarrow z'_1 \ldots\ldots z'_m \quad \text{iff}$$

$$z_1 z_\sigma(0)_s \ldots z_m z_\sigma(0)_s \Rightarrow z'_1 z_\sigma(0)_s \ldots z'_m z_\sigma(0)_s \text{ iff}$$

$$(20, 8) \quad z_1(0)\, z_\sigma(S_1(z_1))_s \ldots z_m(0) z_\sigma(S_m(z_m))_s \Rightarrow$$

$$\Rightarrow z_1(0)\, z_\sigma(S_1(z'_1))_s \ldots\ldots z_m(0)\, z_\sigma(S_m(z'_m))_s \quad \text{iff}$$

$$(9.i) \quad z_\sigma(S_1(z_1))_s \ldots\ldots z_\sigma(S_m(z_m))_s \Rightarrow z_\sigma(S_1(z'_1))_s \ldots z_\sigma(S_m(z'_m))_s$$

$$\text{iff } (19) \quad \sum_1^m S_i(z_i) \leq \sum_1^m S_i(z'_i)$$

CHAPTER 3: INTERNAL ENERGY

This chapter is an analysis of the conditions which enable us to construct a set of extensive internal energy functions $U_i(z_i)$ for the systems $Z_i \in \mathcal{Z}$

3.1 Def. A set of extensive internal energy functions $f_i(z_i)$ for the systems $Z_i \in \mathcal{Z}$ is a set of real valued functions $f_i(z_i)$ $z_i \in Z_i$, such that

(i) $f_i(z_i') = f_i(z_i'')$ iff $z_i'(u)\ z_i''$

(ii) if $Z \equiv Z_i \times Z_j \times \ldots$ then $f(z) = f(z_i z_j\ ,,,) = f_i(z_i) + f_j(z_j) + \ldots..$

The construction of such a set of extensive energy functions meets a fundamental difficulty in the fact that the relation $(Iu)_i$ or $(u)_i$ or (u) is an equivalence relation, which does not invoke an order in the equivalence classes of "equal internal energy". The method followed in the case of the construction of the extensive entropy functions cannot be used without the introduction of additional properties. There is a simple way to overcome this difficulty. We establish in the energy equivalence classes of the special system which will be used as an energy meter a simple order through the extra requirement that the energy - and entropy equivalence classes of this system are identical. The order of the energy equivalence classes which is described through the relation $\lesssim u)_v$ is then identical with the order established by the relation $\Rightarrow_v$ Thus:

3.2 Def: For all $z_v', z_v'' \in Z$:

(i) $z_v' \lesssim u)\ z_v''$ iff $z_v' \Rightarrow z_v''$

(ii) $z_v' <u)\ z_v''$ iff $z_v' \rightarrow z_v''$

Other requirements of the energy meter will be explained in the detailed exposition below.

Our first objective is to define relations $\lesssim u)$ for compositions of identical energy meters, and to derive a theorem giving properties of the relations $\lesssim u)$, such that this theorem is isomorphic with the metrisation axioms for entropy, however, restricted to compositions of energy meters Z_ν (i.e. the theorem can be obtained by replacing the symbol $\Rightarrow$ by the symbol $\lesssim u)$ and substitution of Z_ν for Z_i, Z_j). This theorem will be derived from the metrization axioms for energy" and the assumed properties of the energymeter.

Then we define a function $U_\nu(z_\nu)$, isomorphic with the function $S_\sigma(z_\sigma)$. The definition of the energy meter Z_ν is such that, by replacing in the definition of the entropy meter the indices s and σ by the indices u and ν and the sign $\rightarrow$ by the sign $<u)$, properties of the energy meter are obtained. This situation allows us to formulate and to prove a theorem stating that $U_\nu(z_\nu)$ behaves as an extensive internal energy function for compositions of identical systems Z_ν, simply by making appropriate substitutions in the isomorphic theorem and proof of the former chapter.

The same procedure leads to a definition of a function $U_i(z_i)$ for all systems $Z_i \in \mathcal{Z}$ and the proof that this function is an extensive internal energy function.

Metrization axioms for energy, the metrization theorem

In the detailed exposition we will give firstly the "metrization axioms for internal energy"

3.3 Ax. For systems $Z_i, Z_j \in \mathcal{L}$ the following holds

(i) $(\forall z_i', z_i'', z_j)\ (z_i'\ (u)\ z_i''$ iff $z_i' z_j\ (u)\ z_i'' z_j)$

(ii) $(\forall z_i^1, \ldots z_i^n)\ (z_i^1 \ldots z_i^n\ (u)$ Permutation $z_i^1 \ldots z_i^n)$

(iii) $(\forall z_i', z_i'')$ (if $(z_i' \ldots z_i')^n\ (u)\ (z_i'' \ldots z_i'')^n$ then $z_i'\ (u)\ z_i'')$

The following theorems are immediate consequences, and will be needed later:

3.4 Th. $(z' \ldots z')^n\ (u)\ (z'' \ldots z'')^n$ iff $z'\ (u)\ z''$

Proof:

$z'\ (u)\ z''$ implies (3.i) $z'z'\ (u)\ z''z'$ and $z'z''\ (u)\ z''z''$.

Thus (3.ii) $z'z'\ (u)\ z''z''$ etc. Axiom (3.iii) offers the other half of the proof. □

3.5 Th. If $z^1 \ldots z^n\ (u)\ (z^* \ldots z^*)^n$ and $z^{n+1} \ldots z^{2n}\ (u)\ (z^{**} \ldots z^{**})^n$ then $z^1 \ldots z^n\ (u)\ z^{n+1} \ldots z^{2n}$ iff $z^*\ (u)\ z^{**}$

The proof is straightforward

3.6 Th. If $z_1 \ldots z_i \ldots z_n\ (u)\ z_1' \ldots z_i' \ldots z_n'$ and $z_i'\ (u)\ z_i''$ then $z_1 \ldots z_i'' \ldots z_n\ (u)\ z_1' \ldots z_i' \ldots z_n'$

Proof:

$z_i\ (u)\ z_i''$ implies (3.i, 3.ii) $z_1 \ldots z_i \ldots z_n\ (u)\ z_1 \ldots z_i'' \ldots z_n$
Thus the transitivity of (u) yields:

$z_1 \ldots z_i'' \ldots z_n\ (u)\ z_1' \ldots z_i' \ldots z_n'$. □

Next we introduce a generalised form of the definition 2 of the relations $\lesssim u)$ and $<u)$

3.7 Def. For systems Z such that $\mathcal{S}(Z) = \mathcal{U}(Z)$:

(i) $z' \ldots z^n \lesssim u)\ z^{n+1} \ldots z^{2n}$ iff $z' \ldots z^n$ (u) $z^* \ldots z^*$
and $z^{n+1} \ldots z^{2n}$ (u) $z^{**} \ldots z^{**}$ and $z^* \Rightarrow z^{**}$

(ii) $z^1 \ldots z^n <u)\ z^{n+1} \ldots z^{2n}$ iff ... idem ... and $z^* \to z^{**}$

where $z^* \ldots z^*$ means $(z^* \ldots z^*)^n$; the brackets and the index n are again omitted because there is never any doubt about the number of states z^* involved.

We derive next a series of theorems which will lead to the important "metrization theorem" that we need to establish the similarity between the theory of this chapter and the former:

3.8 Th. For systems Z such that $\mathcal{S}(Z) = \mathcal{U}(Z)$:
the relation $<u)$ on $(Z \times \ldots \times Z)^n \times (Z \times \ldots \times Z)^n$ is transitive, or
if $z' \ldots z^n <u)\ z^{n+1} \ldots z^{2n}$ and $z^{n+1} \ldots z^{2n} <u)\ z^{2n+1} \ldots z^{3n}$
then $z' \ldots z^n <u)\ z^{2n+1} \ldots z^{3n}$.

Proof:

(7) $z' \ldots z^n <u)\ z^{n+1} \ldots z^{2n}$ iff $z' \ldots z^n$ (u) $z^* \ldots z^*$ and $z^{n+1} \ldots z^{2n}$ (u) $z^{**} \ldots z^{**}$ and $z^* \to z^{**}$;

(7) $z^{n+1} \ldots z^{2n} <u)\ z^{2n+1} \ldots z^{3n}$ iff $z^{n+1} \ldots z^{2n}$ (u) $z^{\dagger} \ldots z^{\dagger}$ and $z^{n+1} \ldots z^{3n}$ (u) $z^{\dagger\dagger} \ldots z^{\dagger\dagger}$ and $z^{\dagger} \to z^{\dagger\dagger}$;

because (u) is transitive: $z^{**} \ldots z^{**}$ (u) $z^{\dagger} \ldots z^{\dagger}$ and thus (3.iii) z^{**} (u) $z^{\dagger}$ and thus $z^{**} \leftrightarrow z^{\dagger}$; $z^* \to z^{**}$ and $z^{**} \leftrightarrow z^{\dagger}$ and $z^{\dagger} \to z^{\dagger\dagger}$ implies $z^* \to z^{\dagger\dagger}$ and thus (5) $z' \ldots z^n <u)\ z^{2n+1} \ldots z^{3n}$. $\square$

3.9 Th. If $z^1 \ldots z^n$ (u) $z^{n+1} \ldots z^{2n}$ and $z^{n+1} \ldots z^{2n} <u) z^{2n+1} \ldots z^{3n}$ then $z^1 \ldots z^n <u) z^{2n+1} \ldots z^{3n}$

Proof:

Suppose $z^{n+1} \ldots z^{2n}$ (u) $z^* \ldots z^*$ and $z^{2n+1} \ldots z^{3n}$ (u) $z^{**} \ldots z^{**}$

Thus (7) $z^* \to z^{**}$. The transitivity of (u) implies $z^1 \ldots z^n$ (u) $z^* \ldots z^*$

Thus (7) $z^1 \ldots z^n <u) z^{2n+1} \ldots z^{3n}$. □

3.10 Th. For systems Z such that $\mathcal{S}(Z) = \mathcal{U}(Z)$ and

$(\forall z^1, \ldots z^n)\ (\exists z^*)\ (z^1 \ldots z^n \text{ (u) } z^* \ldots z^*)$:

$(\forall z^1, \ldots z^{2n})$ (either $z^1 \ldots z^n <u) z^{n+1} \ldots z^{2n}$ or

$z^{n+1} \ldots z^{2n} <u) z^1 \ldots z^n$ or $z^1 \ldots z^n$ (u) $z^{n+1} \ldots z^{2n}$)

Proof:

Suppose $z^1 \ldots z^n$ (u) $z^* \ldots z^*$ and $z^{n+1} \ldots z^{2n}$ (u) $z^{**} \ldots z^{**}$;

because either $z^* \to z^{**}$ or $z^{**} \to z^*$ or $z^{**} \leftrightarrow z^*$ the theorem follows. □

3.11 Th. For systems Z such that $\mathcal{S}(Z) = \mathcal{U}(Z)$:

$(\forall z', z'')\ (z' \ldots z' <u) z'' \ldots z''$ iff $z' <u) z'')$

Proof:

Because (7) $z' \ldots z' <u) z'' \ldots z''$ iff $z' \to z''$

and $z' \to z''$ iff $z' <u) z''$ follows:

$z' \ldots z' <u) z'' \ldots z''$ iff $z' <u) z''$. □

3.12 Th. For systems Z such that (a) $\mathscr{S}(Z) = \mathscr{U}(Z)$,

(b) $(\forall z', \dots, z^m)\ (\exists z^*)\ (z' \dots z^m\ (u)\ z^* \dots z^*)$ and

(c) $(\forall z', \dots, z^{2n}, z)\ (z' \dots z^n <u)\ z^{n+1} \dots z^{2n}$ iff
$z' \dots z^n z <u)\ z^{n+1} \dots z^{2n}\ z)$:

$(\forall z', \dots z^{2n+2m})$ (if $z' \dots z^n <u)\ z^{n+1} \dots z^{2n}$ and
$z^{2n+1} \dots z^{2n+m} <u)\ z^{2n+m+1} \dots z^{2n+2m}$, then
$z' \dots z^n z^{2n+1} \dots z^{2n+m} <u)\ z^{n+1} \dots z^{2n}\ z^{2n+m+1} \dots z^{2n+2m}$)

Proof:

$z' \dots z^n <u)\ z^{n+1} \dots z^{2n}$ implies (c)
$z' \dots z^n\ z^{2n+1} <u)\ z^{n+1} \dots z^{2n} z^{2n+1}$ and this etc. ....
$z' \dots z^n\ z^{2n+1} \dots z^{2n+m} <u)\ z^{n+1} \dots z^{2n} z^{2n+1} \dots z^{2n+m}$;
$z^{2n+1} \dots z^{2n+m} <u)\ z^{2n+m+1} \dots z^{2n+2m}$ implies (c)
$z^{2n+1} \dots z^{2n+m}\ z^{n+1} \dots z^{2n} <u)\ z^{2n+m+1} \dots z^{2n+2m} z^{n+1} \dots z^{2n}$ and
this implies (3.ii.9):
$z^{n+1} \dots z^{2n} z^{2n+1} \dots z^{2n+m} <u)\ z^{n+1} \dots z^{2n}\ z^{2n+m+1} \dots z^{2n+2m}$;
the transitivity of <u) gives finally :
$z' \dots z^n z^{2n+1} \dots z^{2n+m} <u)\ z^{n+1} \dots z^{2n} z^{2n+m+1} \dots z^{2n+2m}$. □

We are now in a position to prove the two theorems we aimed for, because they are similar to theorems of the former chapter.

3.13 Th. For systems $Z \in \mathscr{L}$ such that

(a) $\mathscr{S}(Z) = \mathscr{U}(Z)$ and

(b) $(\forall z', \dots z^n)\ (\exists z^*)\ (z' \dots z^n\ (u)\ z^* \dots z^*)$:

the relation $\lesssim$u) is a simple preorder on $(Z \times \dots \times Z)^n \times (Z \times \dots \times Z)^n$ for all positive integers n.

Proof:

the comparability of $\leq u)$ is implied by theorem 10; the transitivity of $\leq u)$ follows from the transitivity of (u) and theorem 8; the reflexivity of (u) implies the reflexivity of $\leq u)$.

3.14 Th. Metrization theorem:

For systems $Z \in \mathcal{Z}$ such that:

(a) $\mathcal{S}(Z) = \mathcal{U}(Z)$

(b) $(\forall z', \ldots, z^n)\ (\exists z^*)\ (z' \ldots z^n\ (u)\ z^* \ldots z^*)$

(c) $(\forall z', \ldots, z^{2n}, z)\ (z' \ldots z^n < u)\ z^{n+1} \ldots z^{2n}$ iff $z' \ldots z^n\ z < u)\ z^{n+1} \ldots z^{2n}\ z)$

(i) $(\forall z', \ldots, z^{2n}\ z)\ \left(z' \ldots z^n \leq u)\ z^{n+1} \ldots z^{2n}\right.$ iff $\left. z' \ldots z^n\ z \leq u)\ z^{n+1} \ldots z^{2n}\ z\right)$

(ii) $(\forall z', \ldots, z^n)\ \left(z' \ldots z^n \leq u)\ \text{permutation}\ z' \ldots z^n\right)$

(iii) $(\forall z', z'')\ \left(\text{if}\ z' \ldots z' \leq u)\ z'' \ldots z''\ \text{then}\ z' \leq u)\ z''\right)$

Proof:

(i) follows from axiom (3.i) and condition (c);

(ii) is an implication of axiom (3.ii);

(iii) follows from axiom (3.iii) and theorem 11.

The construction of extensive internal energy functions, the energy meter.

We start again with the definition of a function $U_v(z_v)$ for a special system Z_v, which properties will be explained afterwards.

3.15 Def. The function $U_v\ (z_v)$, $z_v \varepsilon Z_v$ is defined as follows:

Choose two states z_v', $z_v'' \varepsilon Z_v$ such that $z_v' \rightarrow z_v''$

$U_v\ (z_v') \equiv 0 \quad U_v\ (z_v'') \equiv 1$; we write $z_v' = z_v(o)_u$,

$z_v'' = z_v(1)_u$.

$U_v\ (z_v) = 1/2$ iff $z_v(o)_u\ z_v(1)_u\ (u)\ z_v\ z_v$ etc.

$U_v\ (z_v) = n$ (positive integer) iff $z_v\ z_v\ (n-2)_u\ (u)\ z_v\ (n-1)_u\ z_v\ (n-1)_u$

$U_v\ (z_v) = -1$ iff $z_v\ z_v(1)_u\ (u)\ z_v(o)_u\ z_v(o)_u$ etc.

$U_v\ (z_v) = r$ (real) iff $\sup \{d:\ z_v(d)_u <u)\ z_v\} = r$

or $\inf \{d:\ z_v <u)\ z_v(d)_u\} = r$ where d is a dyadic.

In the case that there exists a z_v such that $U_v\ (z_v) = n$, and there does not exist a z_v' such that $U_v\ (z_v') = n + 1$, we define $U_v\ (z_v) = n + 1/2$ iff $z_v\ z_v\ (n - 1/2)_u\ (u)\ z_v(n)_u\ z_v(n)_u$.

3.16 Def. An energymeter is a system $Z_v \varepsilon \mathscr{L}$ such that

(i) $\mathscr{S}(Z_v) = \mathcal{U}(Z_v)$

(ii) $(\mathscr{S}, \rightarrow) \simeq (R', <)$

(iii) $(\forall z_v)\ (\forall \mathscr{N}_{\mathcal{J}}(z_v))\ (\exists z_v\ (d')_u,\ z_v\ (d'')_u)$
$\left[z_v\ (d')_u,\ z_v\ (d'')_u\ \varepsilon\ \mathscr{N}_{\mathcal{J}}(z_v) \text{ and } z_v\ (d')_u \rightarrow z_v \rightarrow z_v\ (d'')_u \right]$

("calibration property of the energymeter")

where, again, a neighbourhood $\mathscr{N}_{\mathcal{J}}(z_v)$ is defined as a set $\{z_v : z_v' \rightarrow z_v \rightarrow z_v''\}$

(iv) $(\forall z_v', \ldots, z_v^n)\ (\exists z^*)\ (z_v' \ldots z_v^n\ (u)\ z^* \ldots z^*)$

(v) $(\forall z_v', \ldots z_v^{2n}, z)\ (z_v' \ldots z_v^n <u)\ z_v^{n+1} \ldots z_v^{2n}$ iff
$z_v' \ldots z_v^n\ z <u)\ z_v^{n+1} \ldots z_v^{2n}\ z)$

The property (iii) explains that in every neighbourhood, defined through $\rightarrow$, of a state of the energymeter, are states with dyadic energy values. This property enables us to define for every state z_v a value $U_v\ (z_v)$: the meter can be "calibrated".

We can explain that (iii) and (iv) are independent by similar arguments as we used to demonstrate the independence of the related properties (ii) and (iii) of the entropymeter. The properties (i), (iv) and (v) form the conditions under which the "metrization theorem for energy" (theorem 14) obtains.

Δ

The question arises, whether the energymeter is a highly artificial device for solving mathematical difficulties of the formal theory, or a system which has also an acceptable physical interpretation. We will demonstrate that the energymeter can be interpreted as a calorimeter. The crucial property which leads to this interpretation is $\mathcal{S}(Z_v) = \mathcal{U}(Z_v)$. This property implies that a change in energy, without a change in entropy, is impossible: the system cannot do work, reversibly, in adiabatic isolation, or the internal energy is only a function of the entropy. A system with a one dimensional phase space with U or S as the only independent coordinate fulfills this condition. A calorimeter is essentially such a system: the internal energy is a unique independent variable of the system, the calorimeter is not able to do work.

Measuring energy differences with the energy meter is thus essentially performing calorimetric measurements. The construction of the energy function $U_v(z_v)$ for the calorimeter is closely related to existing methods of calibrating these meters.

A remarkable trait of this interpretation is that the concept of "heat" in the sophisticated form of "internal energy" regains such a central place in thermodynamics. This is, however, understandable if we realise that we aim to construct a thermodynamics independent of mechanics.

Δ As a consequence of definition 7 or definition 2, there exists a similarity mapping from $(\mathcal{S}, \rightarrow)$ onto $(\mathcal{U}, <u))$, where the equivalence classes of equal internal energy of the family $\mathcal{U}$ are ordered through the relation <u), according to the ordering <u) of the states contained in them.

Thus $(\mathcal{S}, \rightarrow) \simeq (\mathcal{U}, <u))$. From this and $(\mathcal{S}, \rightarrow) \simeq (R', <)$ it follows that $(\mathcal{U}, <u)) \simeq (R', <u))$. It further follows from definition 2 that a neighbourhood $\mathcal{N}_s(z_v)$ is also a neighbourhood $\mathcal{N}_u(z_v)$, where the latter is defined as a set $\{z_v^\dagger : z_v' <u)\ z_v^\ddagger <u)\ z_v''\}$ containing z_v. This allows us to rewrite property (iii) of the energy meter and to give an equivalent formulation of definition 16 as follows:

3.16* Def. An energy meter Z_v is a system $Z_v \in \mathcal{Z}$ such that

(i) $\mathcal{S}(Z_v) = \mathcal{U}(Z_v)$
(ii) $(\mathcal{U}, <u)) \simeq (R', <)$
(iii) $(\forall z_v)\ (\forall \mathcal{N}_u (z_v))\ (\exists\, z_v (d')_u, z_v (d'')_u)$

$$\left[z_v (d')_u,\ z_v (d'')_u \in \mathcal{N}_u (z_v) \text{ and } z_v (d')_u <u)\ z_v <u)\ z_v(d'')_u\right]$$

where a neighbourhood $\mathcal{N}_u(z_v)$ is defined as a set

$\{z_v\dagger :\ z_v' <u)\ z_v\dagger <u)\ z_v''\}$ containing z_v.

(iv) $(\forall z_v', \ldots, z_v^n)\ (\exists z^*)\ (z_v' \ldots z_v^n\ (u)\ z_v^* \ldots z_v^*)$

(v) $(\forall z_v', \ldots, z_v^{2n}, z)\ \langle z_v' \ldots z_v^n <u)\ z_v^{n+1} \ldots z_v^{2n}$ iff
$z_v' \ldots z_v^n\, z <u)\ z_v^{n+1} \ldots z_v^{2n}\, z\rangle$

A comparison between the metrisation axioms for entropy, the definition of the entropy function $S_\sigma (z_\sigma)$, and of the entropy meter Z_σ at one side and the metrization theorem, the definition of the energy function $U_v(z_v)$, and of the energy meter Z_v at the other side, shows immediately the close formal resemblance. This allows the following theorem:

3.17 Th. For the compositions of identical energy meters Z_v:

$$z_v(r_1)\ z_v(r_2) \ldots z_v(r_n) \lesssim u)\ z_v(r_1')\ z_v(r_2') \ldots z_v(r_n')$$

$$\text{iff} \quad \sum_1^n r_k = \sum_1^n r_k'$$

Proof:

Replace in the proofs of theorems 18 and 19 of chapter 2 $\Rightarrow$ by $\lesssim u)$, the indices σ and s by the indices v and u, and S by U. $\square$ The definition of an energy function $U_i(z_i)$ of an arbitrary system $z_i \in \mathcal{Z}$ does not give special difficulties:

3.18 Def. A function $U_i\ (z_i)$, $z_i \varepsilon Z_i \varepsilon \mathcal{L}$ is constructed as follows:

Choose an energymeter Z_v and a state $z_i' \varepsilon Z_i$

$U_i\ (z_i') \equiv o$; we write $z_i' = z_i(o)_u$

$U_i\ (z_i) = r$ iff $z_i\ z_v(o)_u\ (u)\ z_i(o)_u\ z_v\ (r)_u$

3.19 Def. The system Z_i is U measurable with the energymeter Z_v, if for every $z_i \varepsilon Z_i$ $U_i(z_i) = r$ (a real number).

3.20 Ax. There exists an energymeter $Z_v \varepsilon \mathcal{L}$ such that all systems $Z_i \varepsilon \mathcal{L}$ are U measurable with the energymeter Z_v.

The proof of the next theorem, which states that the set of functions $U_i(z_i)$ is a set of extensive energy functions is completely similar to the proof of the extensivity of $S_i(z_i)$:

3.21 Th. For all $Z_i \varepsilon \mathcal{L}$:

$$z_1 \ldots z_n\ (u)\ z_1' \ldots z_n' \text{ iff } \sum_1^n U_k\ (z_k) = \sum_1^n U_k\ (z_k').$$

CHAPTER 4 PHASE SPACE, FUNDAMENTAL EQUATIONS

This chapter introduces the concept of phase space. It will appear in chapter 5 that the phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$, with the internal energy U_i and the deformation coordinates X_{ki} as independent variables, plays an important role. A second important phase space is $[S_i, \ldots X_{ki}, \ldots]$, $k \in N_i$. Phase spaces are R^n spaces, where n is the number of independent variables defining the system Z_i. I do not presuppose that every point of a phase space represents a state $z_i \in Z_i$. Conversely, however, every state $z_i \in Z_i$ is represented by a unique point of phase space. In other words: Z_i is represented by the "occupied" points of a phase space. In the first paragraph of this chapter we will devise a terminology which enables us to describe different ways of "occupation" of phase space.

For a system $Z_i \in \mathcal{Z}$, with phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$, an entropy function $S_i(z_i)$, which can be written as $S(z_i) = f(U_i(z_i), \ldots, X_{ki}(z_i) \ldots)$ is defined. This notation stresses that we are concerned with a set theoretical function with domain $z_i \in Z_i$, still without analytical properties. In traditional thermodynamics the function $S_i(U_i, \ldots, X_{ki}, \ldots)$ defined for all points of the phase space $[U_i, \ldots, X_{ki}, \ldots]$ $k \in N_i$, plays an important role, and generally it is tacitly assumed that this function is differentiable. In the second paragraph I will investigate which conditions have to be fulfilled to guarantee the existence of two-sided derivatives $\frac{\partial S_i}{\partial U_i}$ and $\frac{\partial S_i}{\partial X_{ki}}$, $k \in N_i$ at all "occupied" points of phase space. These conditions are of two different kinds. The first kind (23 i,ii) formulates the requirements with respect to the occupation of the phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$. The second kind (23.iii,iv) contains continuity assumptions for the function $S_i(z_i) = f(U_i(z_i), \ldots, X_{ki}(z_i), \ldots)$. This set of sufficient conditions can be investigated from a physical point of view. It appears then that they are physically indistinguishable from stronger conditions (27. i,ii), the latter being much simpler and allowing the development of a more transparent thermodynamic theory.

A second important function is $U_i(z_i) = g(S_i(z_i), \ldots, X_{ki}(z_i), \ldots)$, $z_i \in Z_i$. This function and $S_i(z_i) = f(U_i(z_i), \ldots, X_{ki}(z_i), \ldots)$, are called the fundamental functions.

An investigation of the differentiability conditions of the internal energy function along lines similar to the treatment of the fundamental entropy function leads to similar results (Th. 28).

The fundamental entropy and internal energy functions are closely related. An analysis of their relationships is, however, not undertaken in this axiomatization.

Phase Space, Occupation of phase space

4.1 Def. An extensive phase space of the system Z_i is a subset of the set of extensive variables $\{U_i, S_i, X_{ki}, \ldots\}$ $k \in N$ (N being the set of indices indicating the different kinds of deformation coordinates), such that with every set of values for the variables of this subset there corresponds at most one state z_i.

The following notation will be used for phase spaces:

$[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i \subset N$; $[S_i, \ldots, X_{ki}, \ldots]$, $k \in N_i \subset N$;
$[\ldots, X_{ki}, \ldots]$, $k \in N_i \subset N$; $[U_i]$ etc.

4.2 Def. Systems with phase spaces $[U_i, \ldots, X_{ki}, \ldots]$ and $[S_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$ are called "simple". This term has, however, so many different meanings in different presentations of thermodynamics that the use of it can cause confusion. The property will therefore be mentioned explicitly in the following (e.g. in the extremal principles).

We assume that the systems $Z_i \in \mathcal{Z}_\theta$ are "simple".

4.3 Ax. The systems $Z_i \varepsilon \mathcal{L}_\theta$ have a phase space $[U_i, \ldots, X_{ki}, \ldots]$ $k \varepsilon N_i \subset \dot{N}$.

Δ

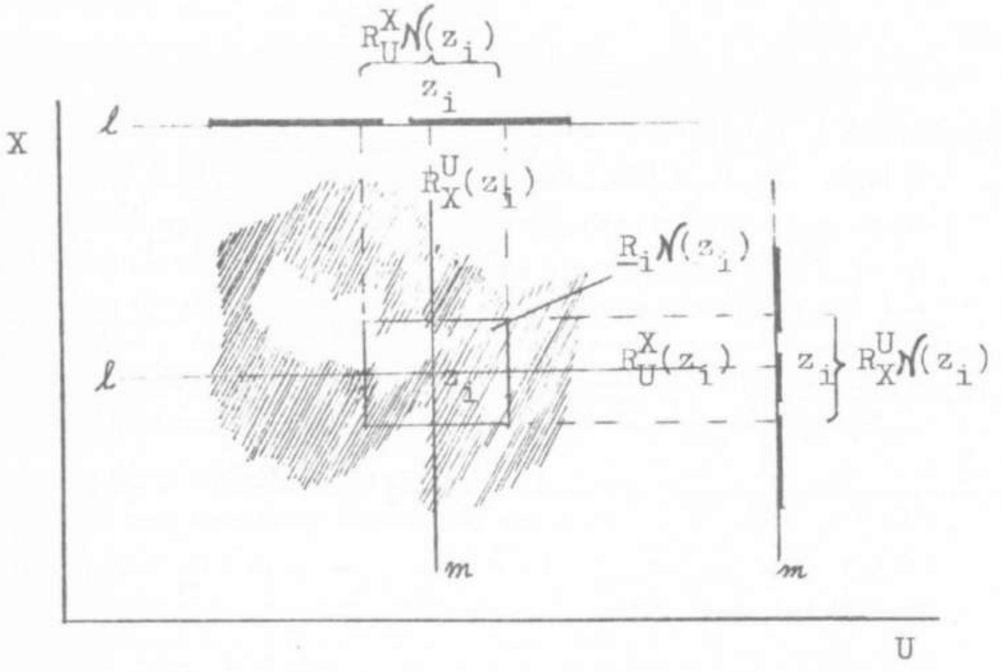
The converse is not true. To show this we consider the so-called "mechanical systems". An important characteristic of these systems is that they do not have a "temperature" or more formally they are not contained in $\mathcal{L}_\theta$. An "internal energy function" can, however, be defined for mechanical systems, which function can be identified with the "purely mechanical" potential energy function. The mechanical energy which is disengaged in a transition from a state of higher potential energy to a state of lower potential energy can be dissipated in a calorimeter (energymeter) and thus measured as a change in internal energy of the mechanical system. An identification of the set of systems with phase space $[S_i, \ldots, X_{ki}, \ldots]$, $k \varepsilon N_i$, and the set $\mathcal{L}_\theta$ is perhaps possible.

Δ An immediate consequence of the definitions of "phase space" and "simple system" is:

4.4 Th. The states z_i of a simple system can be uniquely represented by points (vectors) $\underset{\sim}{r}_i$ in a phase space $\underset{\sim}{R}_i \equiv R_{Ui} \times \ldots \times R_{X_{ki}} \times \ldots.$ $(k \varepsilon N_i)$ and by points (vectors) $\underset{\sim}{r}'_i$ in a phase space $\underset{\sim}{R}'_i \equiv R_{S_i} \times \ldots \times R_{X_{ki}} \times \ldots.$ $(k \varepsilon N_i)$.

Only these two preferred phase spaces will be written in the shortened notation $\underset{\sim}{R}_i$ and $\underset{\sim}{R}'_i$ respectively; other phase spaces will be written as cartesian products of the chosen variables.

The distinction between the points of phase space and the "occupied" points can be expressed by the notation $\underset{\sim}{r}_i$ for point, and $\underset{\sim}{r}_i(z_i)$ for point $\underset{\sim}{r}_i$ occupied by state z_i.



The shaded parts represent existing states z_i.
The bold parts of line ℓ and m represent the subsets $Z^X_U(z_i)$ and $Z^U_X(z_i)$.
The shaded part of the rectangular around z_i represents a neighbourhood $Z_i\mathcal{N}(z_i)$.
The following differences exist between the symbols used in this thesis and those used in my article (1) :

this thesis	(1)	this thesis	(1)
$\underline{R}_i\mathcal{N}(z_i)$	${}^{\underline{R}}{}_{i}\mathcal{N}(z_i)$	$R^X_U\mathcal{N}(z_i)$	${}^{R}{}_{i}\mathcal{N}_U(z_i)$
$R^X_U(z_i)$	$R_U(z_i)$	$R^U_X\mathcal{N}(z_i)$	${}^{R}{}_{i}\mathcal{N}_X(z_i)$
$Z^X_U(z_i)$	$Z_U(z_i)$	$Z_i\mathcal{N}(z_i)$	${}^{Z}{}_{i}\mathcal{N}(z_i)$

$Z^X_U\mathcal{N}(z_i)$ and ${}^{Z}{}_{i}\mathcal{N}_U(z_i)$ are not defined identically.

figure 4.1

The components of the vector $\underset{\sim}{r}_i$ can be written as $U_i(\underset{\sim}{r}_i)$, $X_{ki}(\underset{\sim}{r}_i)$ etc, which has to be distinguished from $U_i(z_i)$, $X_{ki}(z_i)$ etc, the values of U_i, X_{ki}, etc. for state z_i. It is obvious that for $\underset{\sim}{r}_i(z_i) : U_i(\underset{\sim}{r}_i) = U_i(z_i)$ $X_{ki}(\underset{\sim}{r}_i) = X_{ki}(z_i)$ etc.

To explain the properties of the phase spaces of the systems of thermodynamics we introduce a series of terms by formal definitions and clarify their meaning with the help of figure 1, which illustrates different concepts for the simple case of a two dimensional phase space $R_U \times R_X$. In the following I shall often restrict the discussion to the case of the phase spaces $R_U \times R_X$ and $R_S \times R_X$. This will simplify the notation considerably and the extension to the n dimensional case will not present essential difficulties.

A neighbourhood in $\underset{\sim}{R}_i$ of state z_i is defined as a "box" around $\underset{\sim}{r}_i(z_i)$:

4.5 Def. A <u>neighbourhood</u> in $\underset{\sim}{R}_i$ of state z_i : $\underset{\sim}{R}_i\mathcal{N}(z_i) \equiv R\mathcal{N}(z_i) \equiv$ $\{\underset{\sim}{r}_i : U_i(\underset{\sim}{r}_i{}') < U_i(\underset{\sim}{r}_i) < U_i(\underset{\sim}{r}_i{}'') \;\&\; U_i(\underset{\sim}{r}_i{}') < U_i(z_i) < U_i(\underset{\sim}{r}_i{}'')$ $\&\; X_{ki}(\underset{\sim}{r}_i{}') < X_{ki}(\underset{\sim}{r}_i) < X_{ki}(\underset{\sim}{r}_i{}'') \;\&\; X_{ki}(\underset{\sim}{r}_i{}') < X_{ki}(z_i) < X_{ki}(\underset{\sim}{r}_i{}'')$ $k \in N_i\}$ or abbreviated

$$\{\underset{\sim}{r}_i : \underset{\sim}{r}_i{}' < r_i < r_i{}'' \;\&\; \underset{\sim}{r}_i{}' < \underset{\sim}{r}_i(z_i) < \underset{\sim}{r}_i{}''\}.$$

The straight line in $\underset{\sim}{R}_i$ through $\underset{\sim}{r}_i(z_i)$, parallel to the U_i-axes, is called the U subspace in $\underset{\sim}{R}_i$ of z_i:

4.6 Def. The <u>U subspace</u> in $\underset{\sim}{R}_i = R_{Ui} \times R_{Xi}$ is

$$R_U^X(z_i) \equiv \{\underset{\sim}{r}_i : X_i(\underset{\sim}{r}_i) = X_i(z_i)\}.$$

If $\underset{\sim}{R}_i = R_{U_i} \times \ldots R_{Xki} \ldots$, $k \varepsilon N_i$, then we write

$$R_U^{X_k,\, k \varepsilon N_i}(z_i) \equiv \{\underset{\sim}{r}_i : X_{ki}(\underset{\sim}{r}_i) = X_{ki}(z_i), \text{ all } k \varepsilon N_i\}.$$

The set of states $z_i^{\dagger}$ which occupy points of the U subspace in $\underset{\sim}{R}_i$ of z_i is called the U subset in $\underset{\sim}{R}_i$ of z_i.

4.7 Def. The <u>U subset</u> of Z_i in $\underset{\sim}{R}_i = R_{U_i} \times R_{X_i}$ is:

$$Z_U^X(z_i) \equiv \{z_i^{\dagger} : \underset{\sim}{r}_i(z_i^{\dagger}) \varepsilon R_U^X(z_i)\}$$

An open line element containing z_i (or more precise $\underset{\sim}{r}_i(z_i)$) of the U subspace $R_U^X(z_i)$, or otherwise stated the intersection of the line $R_U^X(z_i)$ and a "box" around z_i, without its walls, is called an U neighbourhood in $\underset{\sim}{R}_i = R_{U_i} \times R_{X_i}$ of z_i :

4.8 Def. An <u>U neighbourhood in R_i</u> $= R_{U_i} \times R_{X_i}$ of z_i is :

$$R_U^X \mathcal{N}(z_i) \equiv R \mathcal{N}(z_i) \cap R_U^X(z_i).$$

The set of states $z_i^{\dagger}$, which occupy points between two occupied points of the line $R_U^X(z_i)$ on either side of z_i, is called an U neighbourhood in $\underset{\sim}{R}_i = R_{U_i} \times R_{X_i}$ of state z_i.

4.9 Def. <u>An U neighbourhood in Z_i</u> with phase space $\underset{\sim}{R}_i = R_{U_i} \times R_{X_i}$ of z_i is:

$$Z_U^X \mathcal{N}(z_i) \equiv \{z_i^{\dagger} : U_i(z_i{'}) < U_i(z_i^{\dagger}) < U_i(z_i{''}) \;\&\; U_i(z_i{'}) < U_i(z_i) < U_i(z_i{''}) \;\&\; z_i^{\dagger}, z_i{'}, z_i{''} \varepsilon Z_U^X(z_i)\}$$

In a similar way we define for phase spaces $\underset{\sim}{R}_i = R_{U_i} \times R_{X_i}$ and $\underset{\sim}{R}_i' = R_{S_i} \times R_{X_i}$:

$$\underset{\sim}{R}_i' \mathcal{N}(z_i) \, , \; \overset{U}{R}_X(z_i) \, , \; R_X^S(z_i) \, , \; Z_X^U(z_i), \; Z_X^S(z_i),$$

$$\overset{U}{R}_X \mathcal{N}(z_i) \, , \; R_X^S \quad (z_i), \; Z_X^U \quad (z_i), \; Z_X^S \quad (z_i) \text{ etc.}$$

The next series of concepts describes possible ways of occupation of the phase spaces. Before we introduce these we explain different topologies.

4.10 Def. The topology $(R, \mathfrak{T}_<)$ is the interval topology for the set of reals R, where $\mathfrak{T}$ is an open interval of reals $\{r : r' < r < r''\}$

This definition can be extended to the R^N space without difficulties.

The states ^{z} of a system with a one dimensional phase space R are ordered through the values $r(z) \varepsilon R$ and this ordering gives the possibility of defining another topology, denoted by $(Z, \mathfrak{T}_<)$.

4.11 Def. The topology $(Z, \mathfrak{T}_<)$ is the interval topology for the one dimensional set of states Z, where $\mathfrak{T}$ is an open interval of states $\{z : r(z') < r(z) < r(z'')\}$; the boundaries of these intervals are thus states. The extension of the definition of this topology to systems with more dimensional phase spaces presents certain difficulties, which will be avoided by considering exclusively one dimensional subsets of more dimensional systems Z.

The most complete form of "occupation" of phase space is described as global connectedness in the interval topology $(R, \mathcal{T}_<)$:

4.12 Def. The set $Z_U^X(z)$ is globally connected in the interval topology $(R, \mathcal{T}_<)$ iff $\{U(z^\dagger) : z^\dagger \in Z_U^X(z)\} = R^\dagger$, where $R^\dagger$ is an interval of reals.

Thus the occupied points of the line $R_U^X(z)$ form a connected line element.

4.13 Def. The set Z with phase space $[U, X]$ is globally connected in the topology $(R2, \mathcal{T}_<)$ iff

$$(\forall z) \quad \left[\{U(z^\dagger) : z^\dagger \in Z_U^X(z)\} = R^\dagger\right]$$

$$(\forall z) \quad \left[\, X(z^\dagger) : z^\dagger \in Z_X^U(z)\} = R^{\ddagger}\right]$$

~~This situation is illustrated in figure 1b~~.

A weaker form of "occupation" is "local connectedness in the interval topology $(R, \mathcal{T}_<)$" where we require that every state has a neighbourhood which is globally connected:

4.14 Def. The set $Z_U^X(z)$ is locally connected in the topology $(R_U, \mathcal{T}_<)$ iff $(\forall z)\ (\exists\, Z_U^X\mathcal{N}(z)) \quad [\{U(z^\dagger) : z^\dagger \in Z_U^X\mathcal{N}(z)\} = R^\dagger]$

The extension of the definition to the set Z is similar to the extension of definition 12 to definition 13. ~~Figure 1c gives an illustration of a set Z with phase space $[U, X]$, locally connected in $(R, \mathcal{T}_<)$~~.

Still weaker than "local connectedness in $(R, \mathcal{T}_<)$" is "local connectedness in $(Z, \mathcal{T}_<)$" which is defined as follows:

4.14 Def. The set $Z_U^X(z)$ is <u>locally connected in the topology $(Z_U^X(z), \mathcal{T}_<)$</u> iff:

$$(\forall z)\ (\forall Z_U^X\mathcal{N}(z))\ (\exists z', z'')\ [z', z'' \in Z_U^X\mathcal{N}(z)\ \&\ U(z') < U(z) < U(z'')]$$

This means that there exist no "next" occupied points on either side of z on the line $R_U^X(z)$. A next state, say z^*, implies that $Z_U^X(z)$ can be divided in two subsets $\{z^\dagger : U(z^\dagger) \leq U(z)\}$ and $\{z^{\dagger\dagger} : U(z^{\dagger\dagger}) \geq U(z^*)\}$, which are both closed in $(Z_X^U(z), \mathcal{T}_<)$, and $Z_X^U(z)$ is thus disconnected at z.

The next concept to describe situations in which for every pair of states z' and z'' of the same U subset there exists a state z, such that $U(z') + U(z'') = 2U(z)$. If in an interval of $Z^X_U(z)$, containing z' and z'', this property holds, then the energy values $U(z') + d(U(z'') - U(z'')$ for all dyadics $0 \leq d \leq 1$ are occupied. We will say that the interval of the subspace $R_U^X(z)$ is "dyadically occupied". Again we distinguish the situation in which the property obtains for all states of the subsets $Z_U^X(z)$ etc. and that in which the property obtains for a neighbourhood of every state z.

4.16 Def. The subspace $R_U^X(z)$ is "<u>globally dyadically occupied</u>" iff:

$$(\forall z', z'')\ (\exists z''')\ [z', z'', z''' \in Z_U^X(z)\ \&\ U(z') + U(z'') = 2U(z'')]$$

4.17 Def. The subspace $R_U^X(z)$ is "<u>locally dyadically occupied</u>" iff:

$$(\forall z^\dagger)\ [\text{if } z^\dagger \in Z_U^X(z) \text{ then } (\exists\, Z_U^X\mathcal{N}(z^\dagger))$$

$$(\forall z', z'')\ (\exists\, z''')\ \{z', z'', z''' \in Z_U^X\mathcal{N}(z)\ \&\ U(z') + U(z'') = 2U(z''')\}]$$

We finish this paragraph with the proof of a useful theorem:

4.18 Th. If $Z^X_U(z)$ is locally connected in $(Z^X_U(z), \mathcal{T}_<)$ and $R^X_U(z)$ is locally dyadically occupied, then for all $z^\dagger \in Z^X_U(z)$, there exists a neighbourhood $Z^X_U \mathcal{N}(z\ )$ such that $\{U(z^*) : z^* \in Z^X_U \mathcal{N}(z^\dagger)\}$ is dense in an interval of reals R^* in the interval topology $(R, \mathcal{T}_<)$ (or : $\{U(z^*) : z^* \in Z^X_U(z)\}$ is "locally dense" in $(R, \mathcal{T}_<)$)

Proof:

Consider a neighbourhood $Z^X_U \mathcal{N}(z)$ in $(Z^X_U(z), \mathcal{T}_<)$ for which the following holds:

if $z', z'' \in Z^X_U \mathcal{N}(z)$ then $(\exists\, z''') \{z''' \in Z^X_U(z)$ and $U(z') + U(z'') = 2\, U(z''')$. Suppose $U(z') < U(z) < U(z'')$; the possibility of this supposition is guaranteed by the supposed local connectedness of $Z^X_U(z)$.

Consider the interval of Reals $R^\dagger = \{r^\dagger : U(z') < r^\dagger < U(z'')\}$. All energy values $U(z') + d\, \{U(z'') - U(z')\}$, for all dyadics $0 < d < 1$ are occupied. All reals $r^\dagger \in R^\dagger$ can be written as $r^\dagger = U(z') + r(U(z'') - (U(z')\}$, $0 < r < 1$, and every real $0 < r < 1$ can be defined as the least upper bound of a lower cut of the dyadics $\{d : 0 < d < 1\}$. Thus the set $\{U(z^\dagger) : U(z') < U(z^\dagger) < U(z'')\}$ is dense in the interval $R^\dagger$. $\square$

Δ

Remarks: The distinction between local connectedness in $(Z^X_U(z), \mathcal{T}_<)$ and local connectedness in $(R, \mathcal{T}_<)$ is of a purely mathematical kind and has no physical significance at all. Because every physical measurement contains an uncertainty, the resulting values can never be given as numbers, but are always more or less narrow intervals. The weakest physically controllable assumption about the occupation of phase space is thus local connectedness in $(R, <)$. The same reason makes it impossible to distinguish physically between a dense set and its closure. This situation is of importance in the choice of the axioms for thermodynamics : it is perfectly reasonable not to choose the mathematically weakest assumptions, but to be content with the physically weakest assumptions, which lead to the fundamental thermodynamic equations. This will make it possible as well to avoid the complications of mathematical refinements. My policy will, however, be to try to explain the mathematically weakest assumptions, which we need, and to show afterwards, that the theory can be simplified by using a more physical approach.

Δ The functions $S(U, \ldots, X_{ki} \ldots)$ and $U(S, \ldots, X_k, \ldots)$. Differentiability

The entropy S_i is a set theoretical function defined on the domain Z_i. For a system with phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$, we can write this function as $S_i(z_i) = f(U_i(z_i), \ldots, X_{ki}(z_i), \ldots)$. In this notation the domain of definition is clearly the set of occupied points of phase space.

The function $S_i(U_i, \ldots, X_{ki}, \ldots)$ is considered as distinct from the above function f, in this sense that the domain of definition is the phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$. In view of this distinction, we have to be very careful in the use of the concept of "continuity" : the term "continuous" will never be used without a specification of the topology concerned. Continuity has the usual meaning:

4.19 Def. A function $f: X \mapsto Y$ is $(X, \mathcal{T}_1)$, $(Y, \mathcal{T}_2)$ continuous iff the inverse image $f^{-1}(G)$ of every open set G in $(Y, \mathcal{T}_2)$ is an open set in $(X, \mathcal{T}_1)$, or in other words if for every $G \in \mathcal{T}_2$, $f^{-1}(G) \in \mathcal{T}_1$.

4.20 Def. A function $f: X \rightarrow Y$ is $(X, \mathcal{T}_1)$, $(Y, \mathcal{T}_2)$ continuous at $x \in X$ if, for all open sets H in $(X, \mathcal{T}_1)$, containing x: the inverse image of every open set G in $(Y, \mathcal{T}_2)$ containing $y = f(x)$, contains an open set H^1 in $(X, \mathcal{T}_1)$ which is contained in H.

(or formally $(\forall H, G)\ \ (x \in H \in \mathcal{T}_1\ \ \&\ f(x) \in G \in \mathcal{T}_2)$
$(\exists H^1)\ \ (H^1 \subseteq f(G)\ \ \&\ \ H^1 \in \mathcal{T}_1\ \ \&\ \ H^1 \subseteq H)$).

We also need the concept of "smoothness" of a function at a value:

4.21 Def. A real valued function $y = f(x)$ is smooth at a value $x = a$ iff for every sequence $x_1, x_2, \dots$ such that $\lim_{n \to \infty} x_n = a$

the limit: $\lim_{x_n \to a} \dfrac{f(x_n) - f(a)}{x_n - a}$

exists and is equal for all sequences.

We are now in a position to explain the conditions which guarantee the existence of partial derivatives $\left(\dfrac{\partial S_i}{\partial U_i}\right)_{X_{ki}}$,

$\left(\dfrac{\partial S_i}{\partial X_{ki}}\right)_{U_i,\, X_{\ell i}}$ and $\left(\dfrac{\partial U_i}{\partial X_{ki}}\right)_{S_i,\, X_{\ell i}}$ at the

occupied points $\underset{\sim}{r}_i(z_i)$ of phase space $[U_i, \dots, X_{ki} \dots]$, $k \in N_i$ and $\underset{\sim}{r}'_i\ (z_i)$ of phase space $[S_i, \dots, X_{ki} \dots]$, $k \in N_i$.

4.22 Th. Sufficient conditions for the existence of the derivative $\left(\dfrac{\partial S_i}{\partial U_i}\right)_{X_i}$ at every state $z_i \in Z_i$ are :

For all $z_i \varepsilon Z_i$:

(i) $Z^X_U(z_i)$ is locally connected in $(Z^X_U(z_i), \mathcal{T}_<)$;

(ii) $R^X_U(z_i)$ is locally dyadically occupied ;

(iii) $S_i(z_i{}^{\dagger}) = f(U_i(z_i{}^{\dagger}))$, $z_i{}^{\dagger} \varepsilon Z^X_U(z_i)$, is

$(\{U(z_i{}^{\dagger}) \ : \ z_i{}^{\dagger} \varepsilon Z^X_U(z_i)\}, \mathcal{T}_<)$, $(\{S(z_i{}^{\dagger}) : z_i{}^{\dagger} \varepsilon Z^X_U(z_i)\}, \mathcal{T}_<)$

continuous at z_i ;

(iv) $S_i(z_i{}^{\dagger}) = f(U_i(z_i{}^{\dagger})$, $z_i{}^{\dagger} \varepsilon Z^X_U(z_i)$ is smooth at z_i .

Proof:

Consider a neighbourhood $Z^X_U\mathcal{N}(z)$ for which conditions (i) and (ii) obtain, which implies that $\{U(z^*) : z^* \varepsilon Z^X_U\mathcal{N}(z)\}$ is dense in an interval of reals containing $U(_z)$ (theorem 18). We distinguish two possibilities:

(1) there exists a smallest interval in $\{S(z^{\dagger}) : z^{\dagger} \varepsilon Z^X_U\mathcal{N}(z)\}$ containing $S_{(z)}$. The map of this interval into $\{U(z^{\dagger}) : z^{\dagger} \varepsilon Z^X_U\mathcal{N}(z)\}$ contains an interval containing $U_{(z)}$ and for all states z^* in this interval $S(z^*) = S(z)$. The function f is thus in a neighbourhood $Z^X_U\mathcal{N}_2(z)$ represented by a set of points in $R_U \times R_X$, which determine a unique curve (a straight line), because of the denseness of $\{U(z^*) : z^* \varepsilon Z^X_U\mathcal{N}(z)\}$ in an interval of reals R^*.

(2) there does not exist a smallest interval ~~of~~ in $\{S(z^{\dagger}) : z^{\dagger} \varepsilon Z^X_U\mathcal{N}(z)\}$ containing $S(z)$. This implies the existence of an infinite number of nested intervals containing $S(z)$, in $\{S(z^{\dagger}) : z^{\dagger} \varepsilon Z^X_U \mathcal{N}(z)\}$, which map into an infinite number of nested intervals containing $U(z)$ in $\{U(z^{\dagger}) : z^{\dagger} \varepsilon Z^X_U\mathcal{N}(z)\}$. If this latter nest of intervals approximates an interval of $\{U(z^{\dagger}) : z^{\dagger} \varepsilon Z^X_U\mathcal{N}(z)\}$ then for this interval all the S values are equal to $S(z)$: again we arrive at the existence of a horizontal $S(U)$ curve-piece through z.

If this nest of intervals approximates U(z), then in a neighbourhood $Z_U^X\mathcal{N}(z)$ the function f is represented by a set of points in $R_U \times R_S$ which determine a unique curve piece which may be discontinuous or may have a kink at U(z). Such discontinuities and kinks are, however, excluded by condition (iv).

The existence of a uniquely determined, curve piece through z in R_U x R_S, without discontinuities or kinks at z, implies the existence of the two-sided derivative $\left(\frac{\partial S}{\partial U}\right)_X$ □

It is naturally permissible to replace the conditions (i) - (iv) through stronger conditions. Such a set is the following:

4.24 Th. Sufficient conditions for the existence of the derivative $\left(\frac{\partial S_i}{\partial U_i}\right)_{X_i}$ at every state $z_i \varepsilon Z_i$ are:

For all $z_i \varepsilon Z_i$:

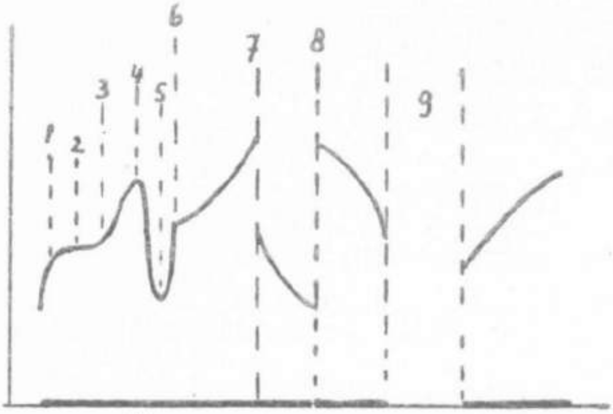
(i) $Z_U^X(z_i)$ is locally connected in $(R, \mathcal{T}_<)$

(ii) $S_i(z_i\dagger) = f(U_i(z_i\dagger)$, $z_i\dagger \varepsilon Z_U^X(z_i)$, is $(R, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous (or otherwise stated: $S_i(U_i)$ is continuous at every $U_i(z_i\dagger) : z_i\dagger \varepsilon Z_U^X(z_i)$)

(iii) $S_i(z_i\dagger) = f(U_i(z_i\dagger))$, $z_i\dagger \varepsilon Z_U^X(z_i)$ is smooth at all z_i (or: $S_i(U_i)$ is smooth at every $z_i\dagger \varepsilon Z_U^X(z_i)$.)

Proof:

It will be immediately clear that condition (i) implies the conditions (i) and (ii) of theorem 23.
In the case of a set $Z_U^X(z_i)$ locally cconnected in $(R, \mathcal{T}_<)$, a neighbourhood $R_U^X\mathcal{N}(z_i)$ is completely occupied: this implies that for a smooth function $S_i(z_i\dagger) = f(U_i(z_i\dagger))$ local



(i) $Z^X_U(z_i)$ is locally connected

(ii) $S_i(U_i)$ is continuous

1 convex upwards

2 horizontal

3 convex downwards

4 maximum

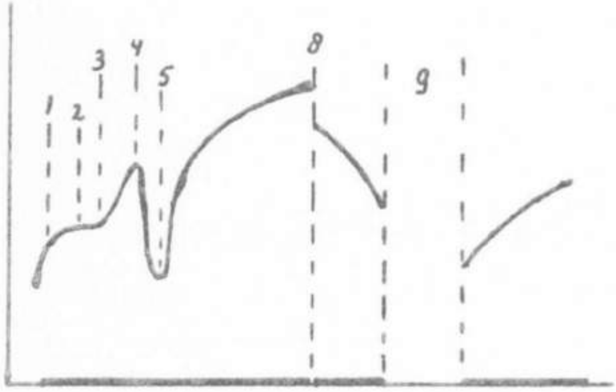
5 minimum

6 kink

7 discontinuity

8 missing state

9 missing interval of states



(iii) $S_i(U_i)$ is smooth

figure 4.2

$(\{U(z^{\dagger}): z^{\dagger} \varepsilon Z_U^X(z_i)\}, \mathcal{T}_<)$ $(\{S(z_i^{\dagger}): z_i^{\dagger} \varepsilon Z_U^X(z_i)\}, \mathcal{T}_<)$ continuity and local $(R, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuity become identical. Thus 4.23 (i) - (iv) may be replacedby 4.24 (i) - (iii)

The conditions 24 (i), (ii) and (iii) can be illustrated geometrically. In figure 2 the thick uninterrupted parts of the U axis represent connected (in $(R, \mathcal{T}_<)$) subsets of $Z_U^X(z_i)$, thus occupied intervals of the one dimensional U space. Interruptions are either unoccupied single points, or closed unoccupied intervals. Condition (ii): continuity of $S(U)_X$ leaves open possibilities for the $S(U)_X$ curve, listed below. Without condition (ii) the existence of a curve is not guaranteed at all. Condition (iii) restricts the possible properties of the $S(U)_X$ curve. Discontinuities and kinks at occupied points are excluded; however, at "missing states" they remain possible. The introduction of the maximum entropy principle in the next chapter will imply a further restriction for the shape of the $S(U)_X$ curve.

Δ

Remark: The distinctions between the weaker conditions 23 (i) - (iv) and the stronger conditions 24 (i) - (iii) for differentiability are of a purely mathematical nature. Physically they are indistinguishable and it is therefore justified to choose the simpler and geometrically more transparent conditions 24 (i) - (iii), when differentiability is required.

Δ

To simplify the terminology in the next chapter we will introduce a few "summarising" concepts:

4.25 Def. The phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \varepsilon N_i$ is locally connected occupied in $(R^n, \mathcal{T}_<)$ at z_i (or $\underset{\sim}{r}(z_i)$) iff

$Z_{U_i}^{X_{ki}}$, $k \varepsilon N_i$ (z_i) is locally connected in $(R, \mathcal{T}_<)$ at z_i and

$Z_{X_{ki}}^{U_i, X_{\ell i}}$ $\ell \varepsilon N_i - k$ (z_i) is locally connected in $(R, \mathcal{T}_<)$ at z_i

for all $k \varepsilon N_i$.

The phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$ is locally connected occupied in $(R^n, \mathcal{T}_<)$ at all $z_i \in Z_i$ means that for every state z_i exist a neighbourhood $\underset{\sim}{R}_i \mathcal{N}(z_i)$ in phase space which is completely occupied. This is a consequence from the foregoing definition.

4.26 Def. The function $\underline{\overline{S_i(z_i^{\dagger}) = f(U_i(z_i^{\dagger}), \ldots, X_{ki}(z_i^{\dagger}), \ldots)}}$ $\underline{\text{is } (R^n, \mathcal{T}_<), \overline{(R, \mathcal{T}_<)} \text{ continuous and smooth at } z_i}$ iff:

$S_i(z_i^{\dagger}) = f_o(U_i(z_i^{\dagger})$, $z_i^{\dagger} \in Z_{U_i}^{X_{ki}}$, $k \in N_i(z_i)$ is
$(R, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ - continuous and smooth at z_i
$S_i(z_i^{\dagger}) = f_k(X_{ki}(z_i^{\dagger})$, $z_i^{\dagger} \in Z_{X_{ki}}^{U_i, X_{\ell i}}$, $\ell \in N_{i-k}\ (z_i)$ is
$(R, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous and smooth at z_i for all $k \in N_i$.

Similar definitions can be formulated for the phase space $[S_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$ and for the function $U_i(z_i^{\dagger}) = g(S_i(z_i^{\dagger}), \ldots, X_{ki}(z_i), \ldots)$. We arrive thus to the final "summarising" results:

4.27 Th. Sufficient conditions for the existence of derivatives

$$\left(\frac{\partial S_i}{\partial U_i}\right)_{X_{ki},\ k \in N_i} \quad \text{and} \quad \left(\frac{\partial S_i}{\partial X_{ki}}\right)_{U_i,\ X_{\ell i},\ \ell \in N_i - k} \quad \text{for all } k,$$

at every state z_i of a system Z_i with phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$, are

(i) the phase space $[U_i, \ldots X_{ki}, \ldots]$, $k \in N_i$ is locally connected in [illegible] $(R^n, \mathcal{T}_<)$ at all $z_i \in Z_i$

(ii) $S_i(z_i) = f(U_i(z_i^{\dagger}), \ldots, X_{ki}(z_i^{\dagger}), \ldots)$ is $(R^n, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous and smooth at all $z_i \in Z_i$.

4.28 Th. Sufficient conditions for the existence of derivatives

$$\left(\frac{\partial U_i}{\partial X_{ki}}\right)_{S_i,\, X_{\ell i},\, \ell \,\varepsilon\, N_i - k} \quad \text{for all } k, \text{ at every}$$

state z_i of a system Z_i with phase space $[S_i, \ldots, X_{ki}, \ldots]$, $k \,\varepsilon\, N_i$ are:

(i) the phase space $[S_i, \ldots X_{ki}, \ldots]$, $k \,\varepsilon\, N_i$, is locally connected occupied in $(R^n, \mathcal{T}_<)$ at all $z_i \,\varepsilon\, Z_i$.

(ii) $U_i(z_i) = g(S_i(z_i), \ldots, X_{ki}(z_i), \ldots)$ is $(R^n, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous and smooth at all $z_i \,\varepsilon\, Z_i$.

CHAPTER 5: EXTREMAL PRINCIPLES, ABSOLUTE TEMPERATURE, ABSOLUTE FORCES

The aim of this chapter is to show the central part which the Gibbs maximum entropy principle and the "mechanical" minimum internal energy principle can play in the development of thermodynamics.

In the first paragraph a precise set theoretical statement of the above principles will be given, which avoids the ambiguities of traditional presentations.

The "traditional" maximum entropy principle reads as follows:

In an equilibrium state the entropy of a system is maximal compared with the entropy of neighbouring (non-equilibrium) states, which are obtained by allowed variations under certain constraints: $\delta S_{U, \ldots X_k, \ldots} \lesssim o$. There exists in traditional thermodynamics no counterpart for the second principle. The Gibbs minimum energy principle, $\delta U_{S, \ldots X_k, \ldots} \gtrsim o$, cannot be considered as such because it restricts itself to thermally homogeneous equilibrium states. Different extremal principles are explained without investigating in their precise relationships. The second paragraph considers the conditions which enable us to demonstrate that the hyper surface $S_i(U_i, \ldots, X_{ki}, \ldots)$ is convex upwards at every "occupied" point of the phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$ of a system Z_i, as a consequence of the maximum entropy principle, and that the hyper surface $U_i(S_i, \ldots, X_{ki}, \ldots)$ is convex downwards at every z_i, as a consequence of the "mechanical" minimum energy principle. Again we find sets of sufficient conditions (in theorems 5, 6 and 7) which are "physically indistinguishable" from stronger but much simpler conditions (in theorems 8 and 9). The results of the second paragraph lead to the most important consequences of the extremal principles, namely that:

$z_i z_j \in C_\theta$ (i.e. z_i and z_j are in thermal equilibrium) iff

$$\frac{\partial S_i(z_i)}{\partial U_i} = \frac{\partial S_j(z_j)}{\partial U_j} .$$

$z_i \; z_j \; \varepsilon \; C_{\phi k}$ (i.e., z_i and z_j are in k^{th} force equilibrium) iff

$$\frac{\partial U_i(z_i)}{\partial X_{ki}} = \frac{\partial U_j(z_j)}{\partial X_{kj}}$$

An absolute temperature function (or better a negative reciprocal absolute temperature function) $N(z_i) \equiv -\frac{1}{T(z_i)} \equiv \frac{\partial S_i(z_i)}{\partial U_i}$

and absolute generalised force functions

$$F_k \equiv \frac{\partial U_i(z_i)}{\partial X_{ki}}$$ can now be defined.

These results can be obtained if certain conditions are imposed on the systems concerned. These conditions can again be chosen in different ways. In theorem 10 I list a weak set of conditions, in theorems 11 and 12 a strengthened set which are identical with the strong version used for the proof of the convexity statements. The two sets are again physically indistinguishable.

Finally, we investigate the conditions which lead to the fundamental equation $T \; d \; S_i = d \; U_i - \sum_{k \, \varepsilon \, N_i} F_k \; d \; X_{ki}$.

§1. Extremal principles

The two principles which form the core of this chapter are the Gibbs maximal entropy principle and the "mechanical" minimum energy principle.

5.1 Ax. Maximum entropy principle

For systems Z_i, $Z_j \varepsilon \mathcal{Z}_\Theta$,
Z_i with phase space $\underline{R}_i = R_{U_i} \times \ldots R_{X_{pi}} \times \ldots$, $p \varepsilon N_i \subset N$
Z_j with phase space $\underline{R}_j = R_{U_j} \times \ldots R_{X_{qj}} \times \ldots$, $q \varepsilon N_j \subset N$
(where N is a set of integers indicating the different kinds of deformation coordinates): $z_i\ z_j \varepsilon C_\Theta \cap \ldots \cap C_{\phi_k} \cap \ldots$,
$k \varepsilon K \subset N_i \cap N_j$, iff
there exists a neighbourhood $\underline{R}_{ij}\,\mathcal{N}(z_i\ z_j)$ in $\underline{R}_i \times \underline{R}_j$ space.
and there exists a set of paired deformation coordinates
X_{ki}, X_{kj}, $k \varepsilon K \subset N_i \cap N_j$ associated with the connections C_{ϕ_k},
such that, for every $z_i^\dagger\ z_j^\dagger$ contained in the neighbourhood
$\underline{R}_{ij}\,\mathcal{N}(z_i\ z_j)$,
if $U_i(z_i^\dagger) + U_j(z_j^\dagger) = U_i(z_i) + U_j(z_j)$,
and $X_{ki}(z_i^\dagger) + X_{kj}(z_j^\dagger) = X_{ki}(z_i) + X_{kj}(z_j)$ for all $k \varepsilon K$,
and $X_{\ell i}(z_i^\dagger) = X_{\ell i}(z_i)$, for all $\ell \varepsilon N_i - K$,
and $X_{mj}(z_i^\dagger) = X_{mj}(z_j)$, for all $m \varepsilon N_j - K$,
then $S_i(z_i^\dagger) + S_j(z_j^\dagger) \leq S_i(z_i) + S_j(z_j)$.

5.2 Ax. Minimum energy principle

For systems $Z_i, Z_j \in \bigcap_{k \in K} \mathcal{Z}_{\phi_k}$

Z_i with phase space $\underset{\sim}{R}'_i = R_{Si} \times \ldots \times R_{X_{pi}} \times \ldots$, $p \in N_i \subset N$,

Z_j with phase space $\underset{\sim}{R}'_j = R_{S_j} \times \ldots \times R_{X_{qj}} \times \ldots$, $q \in N_i \subset N$:

$z_i\, z_j \in \ldots . \cap\ C_{\phi_k} \cap \ldots\ k \in K\ \subset N_i \cap N_j$, iff

there exists a neighbourhood $\underset{\sim}{R}'_{ij}\ \mathcal{N}(z_i z_j)$ in $\underset{\sim}{R}'_i \times \underset{\sim}{R}'_j$ space, and there exist sets of paired deformation coordinates X_{ki}, X_{kj}, $k \in K \subset N_i \cap N_j$ associated with the connections $C_{\phi k}$, such that, for every $z_i^{\dagger}\, z_j^{\dagger}$ contained in the neighbourhood $\underset{\sim}{R}'_{ij}\ \mathcal{N}(z_i z_j)$,

if $S_i(z_i^{\dagger}) = S_i(z_i)$ and $S_j(z_j^{\dagger}) = S_j(z_j)$,

and $X_{ki}(z_i^{\dagger}) + X_{kj}(z_j^{\dagger}) = X_{ki}(z_i) + X_{kj}(z_j)$, for all $k \in K$

and $X_{\ell i}(z_i^{\dagger}) = X_{\ell i}(z_i)$ for all $\ell \in N_i - K$,

and $X_{mj}(z_j^{\dagger}) = X_{mj}(z_j)$ for all $m \in N_j - K$,

then $U_i(z_i^{\dagger}) + U_j(z_j^{\dagger}) \geqslant U_i(z_i) + U_j(z_j)$.

The minimum energy principle has to be distinguished from the "Gibbs minimum energy principle" which says that an equilibrium state $z_i z_j$ of two systems which are thermally and force connected is a state of minimum internal energy compared with neighbouring (non-) equilibrium states, under certain constraints, or more precisely:

5.3 Gibbs minimum energy principle

For systems $Z_i, Z_j \in \mathcal{Z}_{\Theta} \cap \mathcal{Z}_{\phi k}$

Z_i with phase space $\underset{\sim}{R}_i = R_{Ui} \times \cdots \times R_{Xpi} \times \cdots$, $p \in N_i \subset N$

Z_j with phase space $\underset{\sim}{R}_j = R_{Uj} \times \cdots \times R_{Xqj} \times \cdots$, $q \in N_j \subset N$:

$z_i z_j \in C_{\Theta} \cap \ldots \cap\ C_{\phi k} \cap \ldots$, $k \in K \subset N_i \subset N$, iff

there exists a neighbourhood $\underset{\sim}{R}_i \times \underset{\sim}{R}_j\ \mathcal{N}(z_i z_j)$, and

there exist pairs of deformation coordinates X_{ki}, X_{kj}, $k \in K \subset N_i \cap N_j$, associated with the connections $C_{\phi k}$, such that

for every $z_i^{\dagger}\, z_j^{\dagger}$ contained in $\underset{\sim}{R}_i \times \underset{\sim}{R}_j\, \mathcal{N}(z_i z_j)$:
if $S_i(z_i^{\dagger}) + S_j(z_j^{\dagger}) = S_i(z_i) + S_j(z_j)$
and $X_{ki}(z_i^{\dagger}) + X_{kj}(z_j^{\dagger}) = X_{ki}(z_i) + X_{kj}(z_j)$ for all $k \in K$
and $X_{\ell i}(z_i^{\dagger}) = X_{\ell i}(z_i)$, for all $\ell \in N_i - K$
and $X_{mj}(z_j^{\dagger}) = X_{mj}(z_j)$ for all $m \in N_j - K$,
then $U_i(z_i^{\dagger}) + U_j(z_j^{\dagger}) \geq U_i(z_i) + U_j(z_j)$

The "mechanical" minimum energy principle applies in the above formulation to systems for which an entropy function is defined. It may be that this entropy function S_i is trivial in the sense that all states are reversibly adiabatically accessible and thus all states $z_i \in Z_i$ have the same entropy.
Let us call such systems "purely mechanical" and collect these systems in a set $\mathcal{Z}_M$.
For such systems the minimum energy principle reduced to the following statement:

5.4 Th. For systems $Z_i, Z_j \in \bigcap_{k \in K} \mathcal{Z}_{\phi k} \cap \mathcal{Z}_M$
Z_i with phase space $\underset{\sim}{R}_i = \ldots \times R_{X_{pi}} \times \ldots, p \in N_i \subset N$
Z_j with phase space $\underset{\sim}{R}_j = \ldots \times R_{X_{qj}} \times \ldots, q \in N_j \subset N$:
for every $z_i z_j \in \ldots \quad C_{\phi k} \quad \ldots, k \in K \subset N_i \cap N_j$,
there exists a neighbourhood $\underset{\sim}{R}_{ij}\, \mathcal{N}(z_i z_j)$ in $\underset{\sim}{R}_i \times R_j$ space,
and there exist sets of paired deformation coordinates
$X_{ki}, X_{kj}, k \in K \subset N_i \cap N_j$, associated with the connections $C_{\phi k}$,
such that for every $z_i^{\dagger}\, z_j^{\dagger}$ contained in the neighbourhood
$\underset{\sim}{R}_{ij}\, \mathcal{N}(z_i z_j)$,

if $X_{ki}(z_i\dagger) + X_{kj}(z_j\dagger) = X_{ki}(z_i) + X_{kj}(z_j)$,
and $X_{\ell i}(z_i{}^{\dagger}) = X_{\ell i}(z_i)$ for all $\ell \in N_i - \mathbf{K}$,
and $X_{mj}(z_j\dagger) = X_{mj}(z_j)$ for all $m \in N_j - K$,
then $U_i(z_i\dagger) + U_j(z_j\dagger) \geq U_i(z_i) + U_j(z_j)$.

Δ

It will be clear that this theorem can be interpreted as the minimum energy principle of mechanics. It is thus a special case of the more general minimum energy principle (ax.2). This is of interest for an investigation of the relationships of thermodynamics and mechanics.

It is not my intention to give here an analysis of these relationships, or of the inter-relations of the four principles. formulated above. Only the first two will be used, and these are independent in the framework of this axiomatisation.

Δ

Convexity

I shall now develop the main consequences of the maximum entropy and minimum energy principles. I give the theorems in their simplest form to avoid complicated notations:

5.5 Th. For a system $Z \varepsilon \mathcal{Z}_\theta$, with phase space $[U, X]$ and such that the conditions 4.23 i-iv hold:

the curve $S(U)_X$ describing uniquely the function

$S(z^\dagger) = f(U(z^\dagger))$, $z^\dagger \varepsilon Z_u^x(z)$, is convex upwards

at every occupied point $U(z)$.

Proof:

We have already shown (4.18, 4.23) that the conditions 23. i-iv imply that for all $z \varepsilon Z$ a neighbourhood $Z_u^x \mathcal{N}(z)$ is dense in a neighbourhood $R_U^x \mathcal{N}(z)$, and that the occupied points of the latter form the domain of a function which is represented as a subset of a unique curve $S(U)$ which is continuous and smooth at every $U(z)$. We say that "the curve $S(U)$ describes uniquely the function $S(z^\dagger) = f\ (U(z^\dagger))$, $z \varepsilon$ $Z_u^x(z)$". There exists thus a neighbourhood $R_u^x \mathcal{N}'(z) \subset R_u^x \mathcal{N}(z)$ such that the $S(U)$ curve in this neighbourhood is continuous and smooth at all points of U space. The convexity upwards of this curve at z follows immediately from the maximum entropy principle. ☐

A similar theorem, and proof, can be obtained by carrying out the substitutions $U \to X$ and $X \to U$ in the above theorem and proof:

5.6 Th. For a system $Z \in \mathcal{Z}_\theta \cap \mathcal{Z}_\phi$ with phase space $[U, X]$ and such that conditions 4.23 i-iv (X substituted for U and U for X) hold: the curve $S(X)_u$ describing uniquely the function $S(z^\dagger) = f(X(z^\dagger))$, $z^\dagger \in Z_x^u(z)$, is convex upwards at every occupied point $U(z)$.

The minimum energy principle (ax. 5.2) gives the following theorem:

5.7 Th. For a system $Z \in \mathcal{Z}_\theta$ with phase space $[S, X]$ and such that conditions 4.23 i-iv (X substituted for U and S for X) hold: the curve $U(X)_s$ describing uniquely the function $U(z^\dagger) = g(X(z^\dagger))$, $z^\dagger \in Z_x^s(z)$, is convex downwards at every occupied point $X(z)$.

We may replace the conditions 4.23 i-iv by the stronger conditions 4.27 i-ii which leads to the simplified and physically equally weak version:

5.8 Th. For a system $Z_i \in \mathcal{Z}_\theta \cap \ldots \cap \mathcal{Z}_{\phi k} \cap$ ..... with phase space $[U_i, \ldots., X_{ki}, \ldots.]$, $k \in N_i$ and such that:

(i) this phase space is locally connected occupied in $(R^n, \mathcal{J}_<)$ at all $z_i \in Z_i$

(ii) $S_i(z_i) = f(U_i(z_i), \ldots, X_{ki}(z_i), \ldots)$ is $(R^n \mathcal{J}_<)$, $(R, \mathcal{J}_<)$ continuous and smooth at all $z_i \in Z_i$; the hypersurface S_i $S_i(U_i, \ldots, X_{ki}, \ldots)$ in the space $\underset{\sim}{R}_i \times R_{S_i}$ is convex upwards at every occupied point of $\underset{\sim}{R}_i$.

Similarly we arrive at:

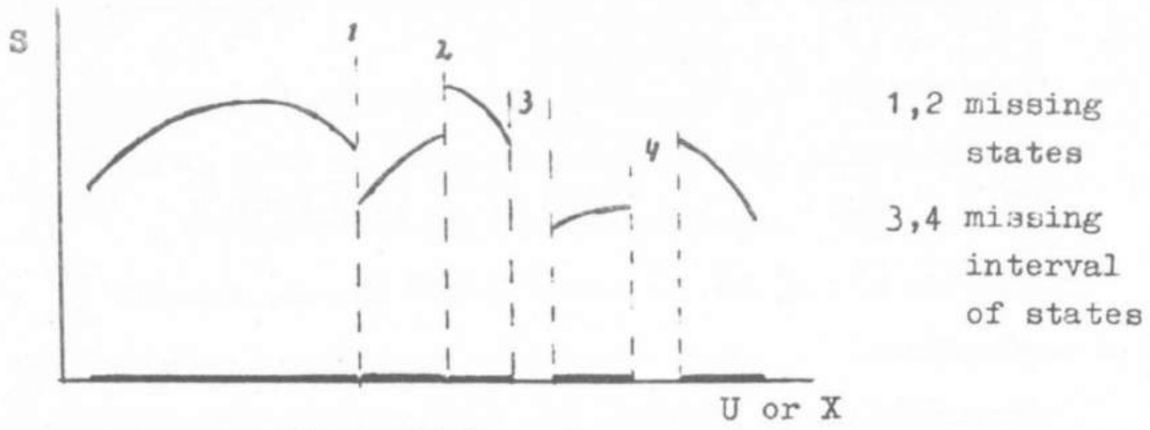


figure 5.1

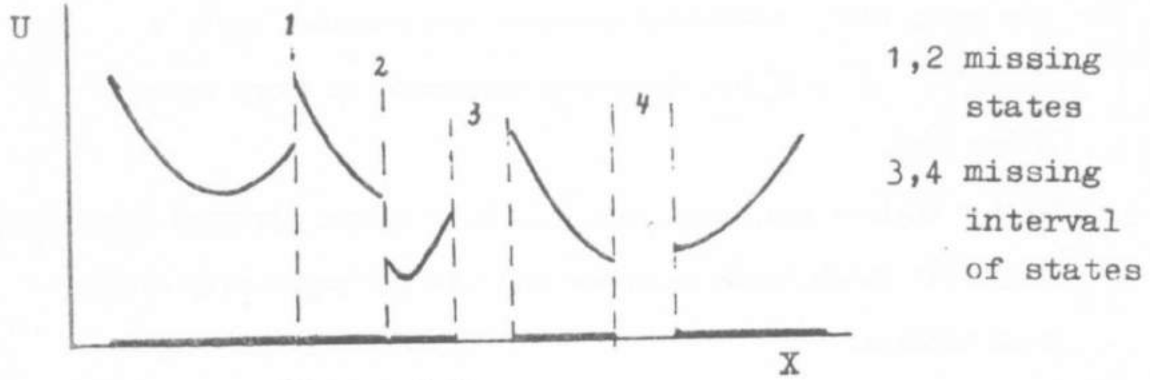


figure 5.2

5.9 Th. For a system $Z_i \in \bigcap_{k \in N_i} \mathcal{L}_{\phi k}$, with phase space $[S_i, \ldots, X_{ki} \ldots]$ $k \in N_i$ and such that:

(i) this phase space is locally connected occupied in $(R^n, \mathcal{T}_<)$ at all $z_i \in Z_i$,

(ii) $U_i(z_i) = g(S_i(z_i), \ldots, X_{ki}(z_i), \ldots)$ is $(R^n \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous and smooth at all $z_i \in Z_i$:

the hyper surface $U_i(S_i, \ldots, X_{ki}, \ldots)$ in the space $\underset{\sim}{R}_i' \times R_{Ui}$ is convex downwards at every occupied point of $\underset{\sim}{R}_i'$.

The above two theorems are illustrated in the following figures (fig. 5.1, 5.2) which give S_i, U_i and S_i, X_{ki} intersections of the space $\underset{\sim}{R}_i \times R_{S_i}$ and an U_i, X_{ki} intersection of the space $\underset{\sim}{R}_i \times R_{Ui}$ (the spaces $\underset{\sim}{R}_i \times R_{S_i}$ and $\underset{\sim}{R}_i' \times R_{Ui}$ are identical!)

§2 Absolute temperature, absolute forces

Next I shall prove that, under certain conditions, for two systems Z_i and Z_j: $z_i\, z_j \in C_\theta$ (i.e., the systems are in thermal equilibrium) iff

$$\frac{\partial S_i(z_i)}{\partial U_i} = \frac{\partial S_j(z_j)}{\partial U_j}$$

Again we can choose weaker and stronger conditions, making the theory more or less complicated. I will prove the theorem starting from a weak set of conditions, and then show that a stronger set is mathematically more transparent and convenient and physically more acceptable. Again also we give the theorem for the case of systems with simple phase spaces to avoid non-essential complications in notation:

5.10 Th. For systems $Z_i, Z_j \in \mathcal{Z}_\theta$, with phase spaces $[U_i, X_{ki}]$ and $[U_j, X_{\ell j}]$ and such that

(i) - (iv): the conditions 4.23 i - iv

(v): $(\forall z_i, z_j)$ if $z_i\, z_j \in C_\theta$ then

$$(\forall Z^{X_k}_{U}\mathcal{N}(z_i),\ Z^{x_\ell}_{U}\mathcal{N}(z_j))\ (\exists z_i', z_i'', z_j', z_j'')$$
$$(z_i', z_i'' \in Z^{x_k}_{U}\mathcal{N}(z_i)\ \&\ z_j'\, z_j'' \in Z^{x_\ell}_{U}\mathcal{N}(z_j)\ \&$$
$$U_i(z_i') < U_i(z_i) < U_i(z_i'')\ \&\ U_j(z_j') < U_j(z_j) < U_j(z_j'')$$
$$\&\ U_i(z_i') + U_j(z_j'') = U_i(z_i'') + U_j(z_j') =$$
$$U_i(z_i) + U_j(z_j)\ )$$

$$z_i\, z_j \in C_\theta \text{ iff } \left(\frac{\partial S_i(z_i)}{\partial U_i}\right)_{X_{ki}} = \left(\frac{\partial S_j(z_j)}{\partial U_j}\right)_{X_{\ell j}}$$

Proof:

Conditions i - iv - together with the maximum entropy principle - imply that the function $S_i(z_i) = f(U_i(z_i))$ for a domain Z [illegible] $_{U}\mathcal{N}(z_i)$ is described by a continuous and smooth curve i through z_i in the plane

$R_U \times R_S$ which is convex upwards (see figure 5.3). A similar conclusion follows for the function $S_j(z_j) = f(U_j(z_j))$

Suppose $z_i\ z_j \in C_\theta$ and $$\frac{\partial S_i(z_i)}{\partial U_i} > \frac{\partial S_j(z_j)}{\partial U_j} > 0$$

Condition v guarantees that every neighbourhood $Z_U^{\times k} \mathcal{N}(z_i)$ and $Z_U^{\times \ell} \mathcal{N}(z_j)$ contains states z_i', z_j' such that
$U_i(z_i') = U_i(z_i) + \Delta U$, $U_j(z_j') = U_j(z_j) - \Delta U$. But the above supposition implies that there exists in every neighbourhood of these kinds states z_i', z_j' such that also: $S_i(z_i') + S_j(z_j') =$
$S_i(z_i) + \Delta S_i + S_j(z_j) + \Delta S_j \approx S_i(z_i) + S_j(z_j) +$

$$\frac{\partial S_i(z_i)}{\partial U_i} \Delta U - \frac{\partial S_j(z_j)}{\partial U_j} \Delta U > S_i(z_i) + S_j(z_j).$$

contradictory to the maximum entropy principle. The only supposition which can be reconciled with the maximum entropy principle is

$z_i\ z_j \in C_\theta$ and $$\frac{\partial S_i(z_i)}{\partial U_i} = \frac{\partial S_j(z_j)}{\partial U_j}$$

Suppose $$\frac{\partial S_i(z_i)}{\partial U_i} = \frac{\partial S_j(z_j)}{\partial U_j}$$

The convexity of the curves i and j has an immediate consequence that for all states z_i', z_j' on these curves, such that
$U_i(z_i') < U_i(z_i)$, $U_i(z_i') + U_j(z_j') =$
$U_i(z_i) + U_j(z_j)$: $S_i(z_i') + S_i(z_i') \leqslant S_i(z_i) + S_j(z_j)$
and thus, according to the maximum entropy principle, $z_i\ z_j \in C_\theta$.
Condition v again guarantees the existence of the states z_i', z_j' in every neighbourhood. □
Similar conditions, imposed on systems Z_i, $Z_j \in \mathcal{Z}_\theta \cap \mathcal{Z}_{\phi k}$ lead to:

$$z_i\ z_j \in C_\theta \cap C_{\phi k} \quad \text{iff} \quad \left(\frac{\partial S_i(z_i)}{\partial X_{ki}}\right)_{U_i} = \left(\frac{\partial S_j(z_j)}{\partial X_{kj}}\right)_{U_j}$$

and similar conditions for systems Z_i, $Z_j \in \mathcal{Z}_{\phi k}$ give

$$z_i\ z_j \in C_{\phi k} \quad \text{iff} \left(\frac{\partial U_i(z_i)}{\partial X_{ki}} \right)_{S_i} = \left(\frac{\partial U_j(z_j)}{\partial X_{kj}} \right)_{S_j}$$

The theory can again be simplified considerably. The condition v of theorem 10 is an implication of condition 4.27 i. We have already seen that the conditions 4.23 i - iv, which appear again in theorem 10 are implied by conditions 4.27 i and ii. We can replace the conditions 10 i-v through the conditions 4.27 i, ii

Δ

This strengthened set of conditions is again physically not distinguishable from the five conditions of theorem 10

Δ

We arrive thus at the following theorems:

5.11 Th. For system Z_i, $Z_j \in \mathcal{Z}_\theta \cap \mathcal{Z}_{\phi k}$, $k \in K \subset N_i \cap N_j$, (K may be empty, then Z_i, $Z_j \in \mathcal{Z}_\theta$), with phase spaces $[U_i, \ldots, X_{pi}, \ldots]$, $p \in N_i$ and $[U_j, \ldots, X_{qj}, \ldots]$, $q \in N_j$, such that

(i) the phase space $[U_i, \ldots, X_{pi}, \ldots]$ is locally connected occupied in $(R^{ni}, \mathcal{T}_<)$ at all $z_i \in Z_i$ and the phase space $[U_j, \ldots, X_{qj}, \ldots]$, $q \in N_j$ is locally connected occupied in $(R^{nj}, \mathcal{T}_<)$ at all $z_j \in Z_j$

(ii) $S_i(z_i) = f_i(U_i(z_i), \ldots, X_{pi}(z_i), \ldots)$ is $(R^{ni}, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous, and smooth at all $z_i \in Z_i$; $S_j(z_j) = f_j(U_j(z_j) \ldots, X_{qj}(z_j), \ldots)$ is $(R^{nj}, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous, and smooth at all $z_j \in Z_j$:

$\underline{a}$ $z_i\ z_j \in C_\theta$ iff $\left(\dfrac{\partial S_i(z_i)}{\partial\ U_i}\right)_{X_{pi}\ \text{all}\ p\ \in\ N_i} = \left(\dfrac{\partial S_j(z_j)}{\partial\ U_j}\right)_{X_{qj}\ \text{all}\ q\ \in\ N_j}$

$\underline{b}$ $z_i\ z_j \in C_\theta \cap C_{\phi k}$ iff

$$\left[\frac{\partial S_i(z_i)}{\partial\ X_{ki}}\right]_{U_i, X_{pi}, \text{all}\ p\ \in\ N_i - k} = \left[\frac{\partial S_j(z_j)}{\partial\ X_{kj}}\right]_{U_j, X_{qj}, \text{all}\ q\ \in\ N_j - k}$$

5.12 Th. For systems Z_i, $Z_j \in \mathcal{Z}_{\phi k}$, $k \in N_i \cap N_j$, with phase space $[S_i, X_{pi}, \ldots]$, $p \in N_i$, and $[S_j, \ldots, X_{qj}, \ldots]$, $q \in N_j$ such that

(i) the phase space $[S_i, \ldots, X_{pi}, \ldots]$, $p \in N_i$ is locally connected occupied in $(R^{ni}, \mathcal{T}_<)$ at all $z_i \in Z_i$, and the phase space $[S_j, \ldots, X_{qj}, \ldots]$, $q \in N_j$ is locally connected occupied in $(R^{nj}, \mathcal{T}_<)$ at all $z_j \in Z_j$

(ii) $U_i(z_i) = g_i(S_i(z_i), \ldots, X_{pi}(z_i), \ldots)$ is $(R^{ni}, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous, and smooth at all $z_i \in Z_i$, $U_j(z_j) = g_j\ (S_j(z_j), \ldots, X_{qi}(z_j), \ldots)$ is $(R^{nj}, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous, and smooth at all $z_j \in Z_j$:

$$z_i\ z_j \in C_{\phi k} \quad \text{iff} \quad \left(\frac{\partial U_i(z_i)}{\partial\ X_{ki}}\right)_{S_i, X_{pi}, \text{all}\ p\ \in\ N_i - k} = \left(\frac{\partial U_j(z_j)}{\partial\ X_{kj}}\right)_{S_j, X_{qj}, \text{all}\ q\ \in\ N_j - k}$$

Next we introduce the concepts "negative reciprocal absolute temperature" and "absolute force".

5.13 Def. The "<u>negative reciprocal absolute temperature function</u>" of a system Z_i, which fulfils the conditions of 4.27, is defined by:

$$N(z_i) \equiv - {}^{1}/_{T(z_i)} \equiv \left(\frac{\partial S_i(z_i)}{\partial U_i}\right)_{X_{pi},\ \text{all } p \in N_i}$$

5.14 Def. The "<u>absolute generalised k^{th} potential function</u>" of a system Z_i, which fulfils the conditions of 4.27, is defined by:

$$F_k(z_i) \equiv \left(\frac{\partial U_i(z_i)}{\partial X_{ki}}\right)_{S_i,\ X_{pi},\ \text{all } p \in N_{i-k}}$$

Theorem 5.11 justifies the name "absolute temperature" for the function $T(z_i)$ The value of the partial derivative $\dfrac{\partial S_i(z_i)}{\partial U_i}$ is uniquely defined through the temperature equivalence class to which z_i belongs, and this equivalence class consists of all states z_i, z_j, z_k of systems Z_i, Z_j, Z_k, ... $\varepsilon\, \mathcal{Z}_\theta$. This justifies us also in writing $T(z_i)$ in place of $T_i(z_i)$. In case of the entropy and internal energy the values $S_i(z_i)$ and $U_i(z_i)$ are dependent on the choice of the state $z_i(o)_s$ and $z_i(o)_u$, which choice is, until now, considered as arbitrary. In case of the deformation coordinates X_{ki} the situation is more complicated. If deformation coordinates are defined via isolations, similar to the entropy and internal energy, then the same applies.

Every system which fulfils the conditions of theorem 4.27 can formally be used as an "absolute thermometer". Weaker conditions will suffice: only the conditions of differentiability of $S_i(z_i) = f(U_i(z_i))$ are necessary.

Absolute temperature values for all states of all systems belonging to $\mathcal{Z}_\theta$ can be defined if the following holds:

5.15 Ax. There exists a set of systems $\mathcal{Z}_T \subset \mathcal{Z}_\theta$ for which the conditions 4.27 i,ii are fulfilled, and such that for all $z_i \varepsilon Z_i \subset \mathcal{Z}_\theta$ there exists a state $z_j \varepsilon Z_j \varepsilon \mathcal{Z}_T$ such that $z_i z_j \varepsilon C_\theta$, i.e., $\mathcal{Z}_\theta$ is "measurable with the set of absolute thermometers $\mathcal{Z}_T$

Again the following weaker condition is sufficient: "there exists a set of states $Z_T = \{\ldots z_i, \ldots z_j, \ldots\}$ such that for all $z_T \varepsilon Z_T$,

$\dfrac{\partial S_T(z_T)}{\partial U_T}$ exists, and for all $z_i \varepsilon Z_i \varepsilon \mathcal{Z}_\theta$ there exists a z_T such that $z_i z_T \varepsilon C_\theta$.

With respect to the absolute potentials F_k and "absolute dynamometers" similar remarks apply, and we make the assumption:

5.16 Ax. There exists a set of systems $\mathcal{Z}_{Fk} \subset \mathcal{Z}_{\phi k}$ such that $\mathcal{Z}_{\phi k}$ is measurable with the set of absolute dynamometers $\mathcal{Z}_{Fk}$.

The last results that we give in this chapter do not require much comment:

5.17 Th. For systems Z_i with phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$ for all states z_i such that

(i) the phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$, is connected occupied in $(R^n, \underset{<}{\mathcal{T}})$ at z_i.

(ii) $S_i(z_i^{\dagger}) = f_o(U_i(z_i^{\dagger}), \; z_i^{\dagger} \in Z_U^{X_k}$, all $k \in N_{i\,(z_i)}$ and

$S_i(z_i^{\dagger}) = f_k(X_{ki}(z_i^{\dagger})), \; z_i^{\dagger} \in Z_X^{U, X_\ell}$, all $\ell \in N_i - k_{(z_i)}$ and

$U_i(z_i^{\dagger}) = g_k(X_{ki}(z_i^{\dagger})), \; z_i^{\dagger} \in Z_{X_k}^{S, X_\ell}$, all $\ell \in N_i - k_{(z_i)}$

are $(R, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$ continuous and smooth at z_i :

(a)
$$\left(\frac{\partial S_i(z_i)}{\partial \; X_{ki}}\right)_{U_i,\, X_{\ell i},\text{ all } \ell \in N_i - k} \left(= - \frac{\partial S_i(z_i)}{\partial \; U_i} \; \frac{\partial U_i(z_i)}{\partial \; X_{ki}} \right) = - \frac{F_k(z_i)}{T(z_i)}$$

(b)
$$d\,S_i(z_i) = \frac{1}{T(z_i)} \; d\,U_i(z_i) - \sum_{k \in N_i} \frac{F_k(z_i)}{T(z_i)} \; d\,X_{ki}(z_i)$$

where $d\,S_i(z_i)$ is a differential at z_i etc.

Defining $d\,Q \equiv T\,d\,S_i$ and $d\,W \equiv \sum_{k \in N_i} F_k \, d\,X_{ki}$ it follows that at z_i: $d\,Q = d\,U_i - d\,W$.

We use the weak conditions (ii) rather than the stronger conditions of theorem 4.27 and 4.28 to allow application of the theorem at states z_i where $S_i(z_i)$ is maximal (see next chapter).

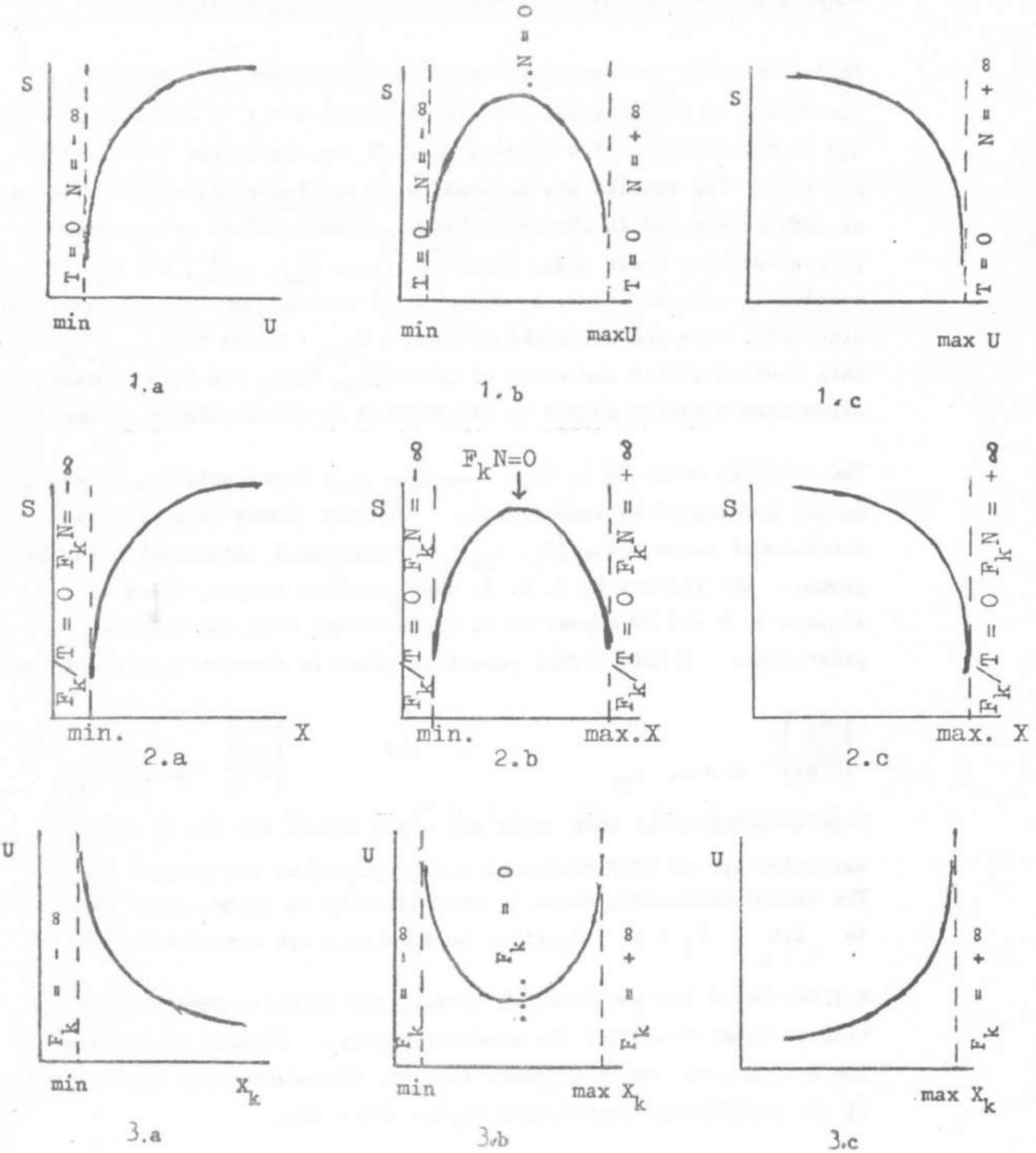
CHAPTER 6: PROPERTIES AT EXTREME TEMPERATURES, THE THIRD LAW

This chapter is an investigation of the properties of systems in the domain of infinite absolute temperature, $T = \pm \infty$, or better $N = o$, and in the domain of low temperatures, $T \rightarrow o$, or better $N \rightarrow - \infty$ and $N \rightarrow + \infty$. The results are derived for a special class of the systems already considered in the last chapter, characterised by the extra properties that their phase space $[U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$, is completely occupied in an n_i-dimensional domain, and that the extremal principles have the strengthened form $\delta\, S_{U,X} < o$ and $\delta\, U_{S_i,\, S_j,\, X} > o$. This implies strict convexity of the $S(U)_X$, $S(X)_U$ and $U(X)_S$ curves, which have possible shapes as illustrated in the figures 1, 2 and 3.

The possible shape of $S_i\ (U_i, \ldots, X_{ki}, \ldots)$ hypersurfaces, in which we are interested is less obvious. For the simple case of a two-dimensional phase space $[U_i, X_{ki}]$ a geometrical illustration can be given. The figures 4, 5, 6, 7, show possible shapes, those of figures 8, 9 and 10 appear to be inconsistent with the extremal principles. A more formal result is given in theorem 3 which states:

$$\left(\frac{\partial S_i}{\partial X_{ki}}\right)_{N = o,\ X_{\ell i}} = o \quad \text{and} \quad \left(\frac{\partial S_i}{\partial U_i}\right)_{N = o,\ X_{\ell i}} = o$$

Next we demonstrate that lower and upper bounds for the U_i and X_{ki} variables are an implication of a lower bound of the entropy S_i. The latter statement, which is introduced as an axiom, leads also to $\lim_{T \rightarrow o} C_v = o$. Positive temperatures are associated with a lower bound for the internal energy, and negative temperatures with an upper bound for the internal energy. Finally we formulate the assumptions, which together with the foregoing cover the content of the traditional formulation of the third law.



figures 6.1-3 : possible shapes of the $S(U)_X$, $S(X)_U$, and $U(X)_S$ curves.

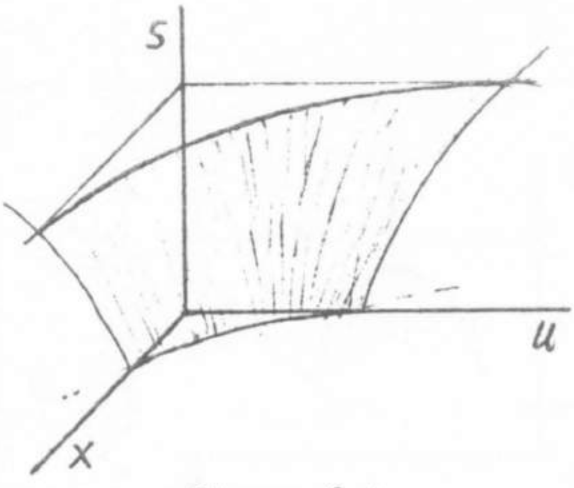


figure 6.4

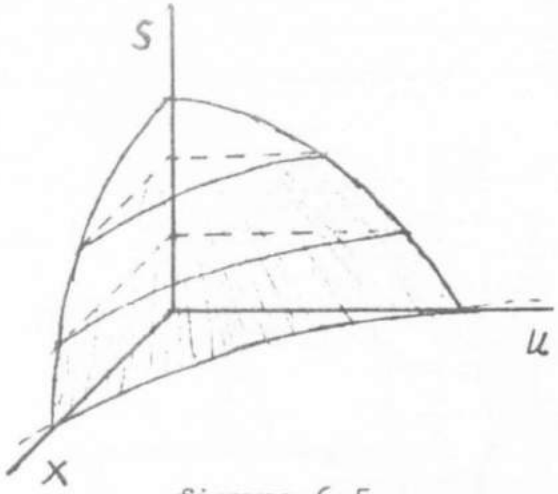


figure 6.5

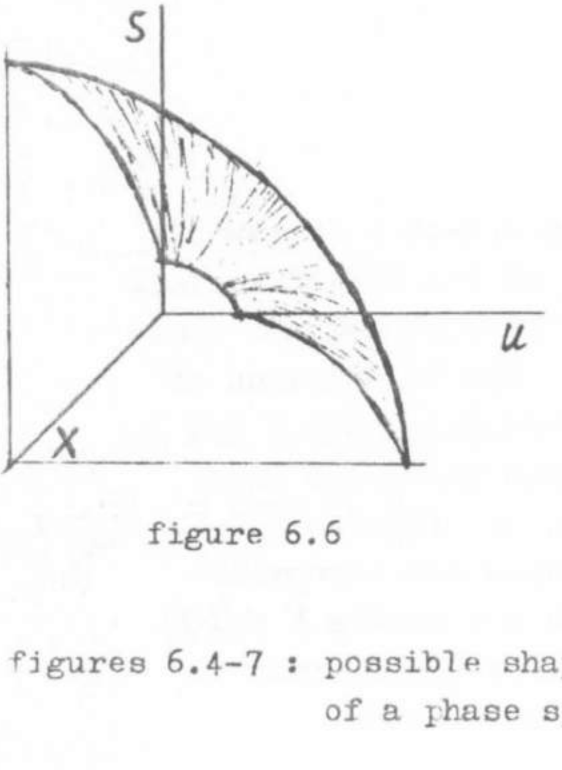


figure 6.6

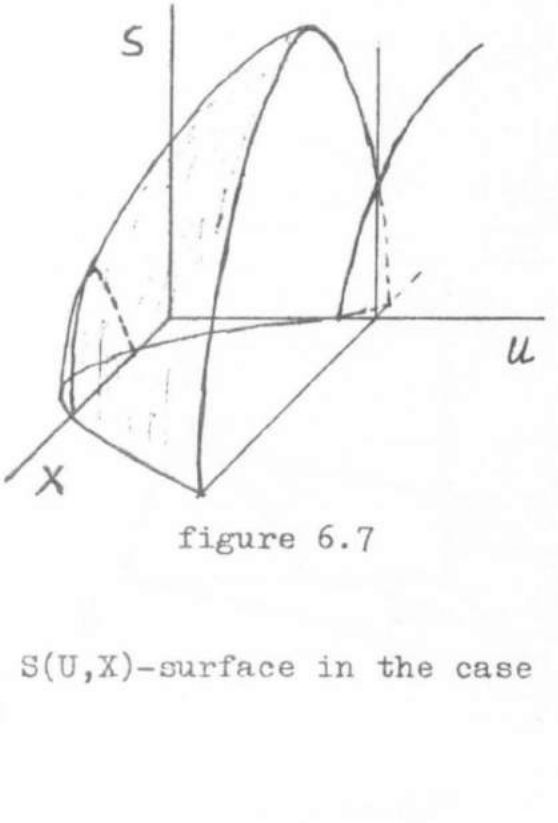


figure 6.7

figures 6.4-7 : possible shapes of the S(U,X)-surface in the case of a phase space [U,X]

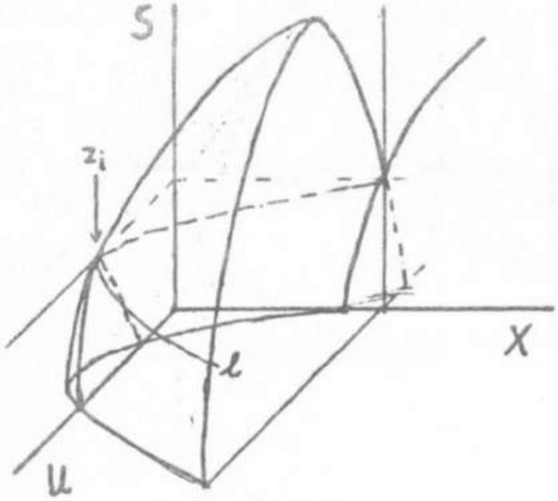


figure 6.8

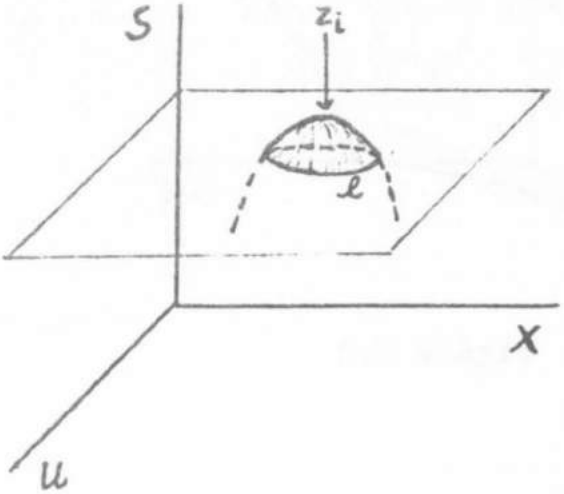


figure 6.9

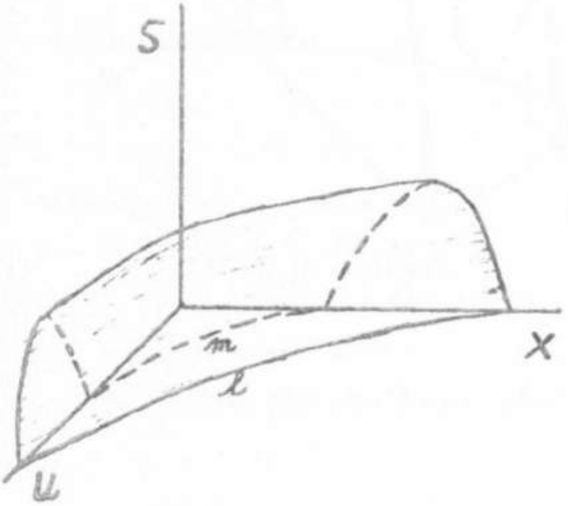


figure 6.10

figures 6.8-10 : impossible shapes of the S(U,X)-surface in the case of a phase space U,X . The orientation of the coordinate axes U and X is chosen different from that in the figures 6.4-7. The shapes are impossible because the curves **ℓ** and ***m*** contradict the minimum energy principle.

6.1 Def. We consider the set of systems $\mathcal{Z}^*_\theta \subset \mathcal{Z}_\theta$, with the following properties:

(i) the systems $Z_i \in \mathcal{Z}^*_\theta$ have a phase space $[U_i, \dots, X_{ki}, \dots]$, $k \in N_i$, which is globally connected occupied in $(R^{n_i}, \mathcal{T}_<)$, i.e. the occupied points fill a connected n_i-dimensional domain of phase space without frontier.

(ii) the function $S_i(z_i) = f\ (U_i(z_i), \dots, X_{ki}(z_i), \dots)$ is $(R^{n_i}, \mathcal{T}_<), (R, \mathcal{T}_<)$ continuous, and smooth at all z_i.

(iii) a strengthened form of the maximum entropy and minimum energy principle holds: namely the unequality signs $\lesssim$ and $\gtrsim$ in axioms 5.1, 5.2 are replaced by $<$ and $>$.

The following theorem follows immediately:

6.2 Th. If $Z_i \in \mathcal{Z}^*_\theta$ then if $Z'_i \subset Z_i$ and if the phase space $[U_i, \dots, X_{ki}, \dots]$, $k \in N_i$ is globally connected occupied by Z'_i, then $Z'_i \in \mathcal{Z}^*_\theta$.

Δ

The physical interpretation of the set $\mathcal{Z}^*_\theta$ requires special attention for the third property. The property (i) is in general tacitly assumed in physical thermodynamics. It is important to notice that $[S_i, \dots, X_{ki}, \dots]$ is <u>not</u> assumed to be necessarily a phase space of $\mathcal{Z}'_\theta$; this leaves open the possibility of the existence of states of maximum entropy. The fact that the occupied domain of phase space has no frontier is mathematically very useful. The intuitive acceptance, in a physical approach of frontiers for the occupied domain clashes with the mathematical usefulness of a refutation of a frontier. The existence and non-existence of a frontier are, however, not physically distinguishable. The property (ii) is generally assumed for all physical systems contained in $\mathcal{Z}_\theta$.

I will not explain how far condition (iii) excludes systems contained in $\mathcal{Z}_\theta$ in its physical interpretation. The condition implies strict convexity upwards of the hypersurface $S(U_i, \ldots, X_{ki}, \ldots)$ and thus

$$\frac{\partial^2 S_i}{\partial\, U_i^{\,2}} < o, \quad \text{etc., i.e.} \quad \frac{1}{T^2}\,\frac{\partial U}{\partial T} > o \quad \text{etc.}$$

Δ

For the systems $Z_i \in \mathcal{Z}^*_\theta$ a few interesting properties can be derived:

6.3 Th. For systems $Z_i \in \mathcal{Z}^*_\theta$: $\left(\frac{\partial S_i(z_i)}{\partial\ X_{ki}}\right)_{N(z_i) = o,\ X_{\ell i},\ \text{all}\ \ell \in N_i - k} = o,$

and $\left(\frac{\partial S_i(z_i)}{\partial\ U_i}\right)_{N(z_i) = o,\ X_{\ell i},\ \text{all}\ \ell \in N_i - k} = o$

<u>Proof:</u>

Consider the subspace $[U_i, X_{ki}]$, containing z_i, or $R^{X_{\ell i},\, \ell \in N_i - k}_{U_i,\, X_{ki}}(z_i)$ (i.e. $\underset{\sim}{r}_i \in [U_i, X_{ki}]$ iff $\underset{\sim}{r}_i \in [U_i, \ldots, X_{ki}, \ldots]$, $k \in N_i$ and $X_{\ell i}(\underset{\sim}{r}_i) = X_{\ell i}(z_i)$). For the states (occupied points) in this subspace the strengthened maximum entropy principle and minimum energy principles hold, and $S_i(U_i, X_{ki})$ is $(R^2, \mathcal{T}_<)$, $(R, \mathcal{T}_<)$-continuous and smooth at all occupied points of the subspace

$$N(z_i) = o \text{ i.e. } \left(\frac{\partial S_i(z_i)}{\partial\ U_i}\right)_{X_{ki}} = o$$

We consider the two possibilities $\left(\frac{\partial S_i(z_i)}{\partial\ X_{ki}}\right)_{U_i} \neq o$ and

$$\left(\frac{\partial S_i(z_i)}{\partial\ X_{ki}}\right)_{U_i} = o$$

If $\left(\frac{\partial S_i(z_i)}{\partial\ X_{ki}}\right)_{U_i} \neq o$ then (see figure 8) the plane of constant entropy through z_i intersects the surface $S_i(U_i, X_{ki})$ in such a way that, at z_i, a straight line parallel to the U_i axis is a tangent of the $U_i(X_{ki})_{S_i}$ curve. One of the parts of this curve which has z_i as its frontier is thus convex upwards and this contradicts the energy minimum principle.

Thus $\left(\frac{\partial S_i(z_i)}{\partial\ X_{ki}}\right)_{U_i} = o$. Because at z_i also $\left(\frac{\partial S_i(z_i)}{\partial\ U_i}\right)_{X_{ki}} = o$ we are left with two possibilities: $S_i(z_i)$ is an absolute maximum (see figure 9) or in every neighbourhood of z_i in $[U_i, X_{ki}]$ are states z_i' such that $S_i(z_i') = S_i(z_i)$, i.e. either there exists a neighbourhood

$R^2_{U_i\,X_{ki}}$ $\mathcal{N}(z_i)$ of states of equal entropy or the states $\{z_i' : S_i(z_i') = S_i(z_i)\}$ form a curve through z_i (see figure 10). If $S_i(z_i)$ is an absolute maximum then (see figure 9) below this state z_i there exists a plane of constant S_i, such that the intersection of this plane with the $S_i(U_i, X_{ki})$ surface is a closed continuous and smooth $U_i(X_{ki})_{S_i}$ curve, which contradicts again the convexity downwards everywhere of the $U_i(X_{ki})_{S_i}$ curves in general. An R^2-neighbourhood of states of equal entropy contradicts the strengthened maximum entropy principle. The only possibility left is thus that the states of maximum entropy in a neighbourhood of z_i form a curve. And, for the states $z_i^{\dagger}$ of this curve:

$$\left(\frac{\partial S_i(z_i^{\dagger})}{\partial \; U_i}\right)_{X_{ki}} = o,$$

it is the isotherm $N = o$. Thus, along this isotherm the entropy is constant:

$$\left(\frac{\partial S_i(z_i)}{\partial \; X_{ki}}\right)_{N = o} = o \quad \text{and} \quad \left(\frac{\partial S_i(z_i)}{\partial \; U_i}\right)_{N = o} = o \qquad \square$$

Next we direct our attention to the domain $\{z_i : N(z_i) \to -\infty\} = \{z_i : T(z_i) \to o\}$ and consider the implication of the following axiom

6.4 Ax. For all $Z_i \in \mathcal{Z}'_{\theta}$: $\{S_i(z_i) : z_i \in Z_i\}$ has a lower bound.

<u>Remark</u>: $\{S_i(z_i) : z_i \in Z_i\}$ cannot have a minimum value because, if there exists such a value, say r, and $S_i(z_i') = r$ then $\{U(z_i) : z_i \in Z^X_U(z_i')\}$ has a minimum contradictory with the

assumption that this is an open interval of reals.

Δ

We have, however, to remember that the last assumption was introduced for reasons of mathematical convenience, and that therefore it is not permitted to give physical significance to the statement that a state with minimum entropy does not exist. The formulation of the axiom leaves both possibilities open.

Δ

6.5 Th. $\{U_i(z_i) : z_i \in Z_i\}$ has a lower bound in the domain $\{z_i : N(z_i) < o\}$ **and an upper bound in the domain** $\{z_i : N(z_i) > o\}$

Proof:

For reasons of convenience we restrict the proof to systems with a two-dimensional phase space $[U_i, X_i]$ **: Consider a curve** $S_i(U_i)_{X_i}$ **, which is continuous, smooth, and convex upwards everywhere.**

(a) Suppose the curve has no maximum, and $\frac{\partial S_i}{\partial U_i} > o$.

Choose a state z_i. $N(z_i) < o$; $T(z_i) > o$

$$U_i(z_i') - U_i(z_i) = \int_{S_i(z_i)}^{S_i(z_i')} T \, d \, S_i$$

A greatest lower bound of $\{U_i(z_i^\dagger) : z_i^\dagger \in Z_{U}^{X}(z_i)\}$ will be approximated, when $S_i(z_i^\dagger)$ approximates the greatest lower bound of $\{S_i(z_i^\dagger) : z_i^\dagger \in Z_{U}^{X}(z_i)\}$.

Say that $\inf \{S(z^\dagger) : z_i{}^\dagger \in Z^X_{\mathcal{U}}(z_i)\} = \rho$.

Then $\inf \{U(z^\dagger) : z_i{}^\dagger \in Z^X_{\mathcal{U}}(z_i)\} - U_i(z_i) = \displaystyle\int_{S_i(z_i)}^{\rho} T \, d\, S_i$

$S_i(z_i)$ is finite for all z_i; ρ is finite; $o < T(z_i{}^\dagger) \leq T(z_i)$ if $\rho < S(z^\dagger) \leq S_i(z_i)$; $U_i(z_i)$ is finite ; thus $\displaystyle\int_{S_i(z_i)}^{\rho} T \, d\, S_i$ is finite and $\inf \{U(z^\dagger) : z^\dagger \in Z^X_{\mathcal{U}}(z_i)\}$ is finite

(b) If the curve $S_i(U_i)_{X_i}$ has no maximum, and $\dfrac{\partial S_i}{\partial U_i} < o$ then a similar proof leads to:
$\sup \{U(z^\dagger) : z^\dagger \in Z^X_{\mathcal{U}}(z_i)\}$ is finite

(c) Suppose $S_i(U_i)_{X_i}$ has a maximum at $z_i{}^*$. Choose a state $z_i{}'$ on the curve such that $U(z_i{}') < U(z_i{}^*)$ and consider $\{z^\dagger : U_i(z_i{}^\dagger) < U_i(z_i{}')\}$. The proof of case (a) leads again to: $\inf \{U(z^\dagger) : U_i(z_i{}^\dagger) < U_i(z_i{}')\}$ is finite. Similarly we find for a state $U_i(z_i{}'') > U(z_i{}^*)$:
$\sup \{U_i(z_i{}^\dagger) : U_i(z_i{}^\dagger) > U_i(z_i{}'')$ is finite. ☐

A similar theorem can be derived with respect to the variables X_{ki} in the domains $\frac{Fk}{T} \to -\infty$ and $\frac{Fk}{T} \to +\infty$

6.6 Th. $\displaystyle\lim_{N \to -\infty} \left(\frac{\partial U_i}{\partial T}\right)_{X_i} = \lim_{T \to o} c_V = o$

<u>**Proof:**</u> Consider a curve $S_i(U_i)_X$.

Call the greatest lower bound of S_i, which is assumed to be approximated if $N \to -\infty$ $(T \to o)$: $S_i(\min)$

$$S(z) - S(\min) =$$

$$\int_o^{T(z)} \frac{dS_i}{\partial U_i} \frac{\partial U_i}{\partial T} \, d\,T = \int_o^{T} \frac{1}{T} \frac{dU_i}{dT} \, d\,T = \int_o^{T} c_V \, d \,\ell n\, T$$

Now $S(z) - S(\min) =$ finite; thus

$\int_o^{T} c_v \, d \,\ell n\, T =$ finite, and this implies that

$$\lim_{T \to o} c_v = o \qquad \Box$$

The content of the "third law" of thermodynamics is covered by the foregoing, and the following statement of axiomatic character:

6.7 Ax.

$$\lim_{\substack{T(z_i) \to o \\ N(z_i) < o}} \left(\frac{\partial S_i(z_i)}{\partial \ U_i}\right)_T = o$$

$$\lim_{\substack{T(z_i) \to o \\ N(z_i) > o}} \left(\frac{\partial S_i}{\partial X_i}\right)_T = o$$

CHAPTER 7 : A CRITICISM OF G. FALK AND H. JUNG. AXIOMATIKDER THERMODYNAMIK, Handbuch der Physik Bd. III/2 p.119-175 Berlin, 1959.

Falk and Jung's article in the Handbuch der Physik is closely related to the axiomatization of the foregoing chapters. The introduction of connections as equivalence relations and isolations as equivalence or order relations of different kinds, the definition of extensive variables entropy and internal energy on the basis of the fundamental properties of the isolations, the use of extremal principles, which relate the connection variables and the isolation variables, the definition of absolute temperature and the derivation of the fundamental thermodynamic equations are common traits. Much of the inspiration of my axiomatisation is derived from Falk and Jung's work.

A critical valuation of Falk and Jung's article meets the difficulty that the proofs of many theorems, which they establish on the basis of the given axioms are not demonstrated explicitly or only sketched. Extra axioms are certainly necessary, and presumably they are omitted because, measured with the standard of rigour of the article, they are certainly self-evident. For instance, the "zerfallende Übergangs relationen" are not explicitly endowed with the reflexive property. This property, added to symmetry and transitivity, makes these relations identical with equivalence relations and this certainly is the intention.

The critical remarks of this chapter are of different kinds. I will investigate whether the theory is consistent and whether the given axioms are sufficient and necessary for the purposes which they serve. Criticism of another kind concerns the physical relevance of the chosen axioms. Following Falk and Jung's article we firstly investigate the metrization axioms of Z‿iff 5 γ and secondly the main chapter of the article "B. Die Struktur der Thermodynamik". The comments follow the division of this chapter : I Entropie und Energie, II Die thermodynamischen Koordinaten.

The Metrization axioms

A comparison between the metrization axioms of Z‿iff 5 γ and the metrization axioms for internal energy of our axiomatization forces itself upon us.

It can be proved that the properties Z‿iff 5 γ (i), (ii), (iii), (vii) and the reflexivity of the energetic isolation implies : $z_i'(u)\ z_i''$ iff $z_i'\ z_j(u)\ z_i''\ z_j$ (axiom 3.3 i.) The relation between Falk and Jungs metrization axioms and the two others in our approach (axiom 3.3 ii, iii)

$$z_i^1 \ldots z_i^n \text{ (u) permutation } z_i^1 \ldots z_i^n \qquad ;$$

$$\text{if } z_i' \ldots z_i' \text{ (u) } z_i'' \ldots z_i'' \text{ then } z_i' \text{ (u) } z_i''$$

is not completely clear. Possibly a construction of rigorous proofs of the theorems of Falk and Jung will show that they are tacitly assumed. If not, then they will follow from the given properties (i), ..., (viii), because they can be considered as an immediate consequence of the additive properties of the internal energy functions U_i whose existence is stated to be an implication of the above properties.

It can further be proved that the properties 5 γ (iv), (v) imply: $(\forall z_i' \, z_i'')$ $(\exists z_i)$ $(z_i' \, z_i'' \, (u) \, z_i \, z_i)$. This latter property is in our axiomatisation restricted to an energymeter, in Falk and Jung's article it is assumed for all systems. Property 5 γ (vi) implies that all energy values $U_i(z_i)$ of a system can be written as dyadics on the basis of a unit energy difference between two arbitrary chosen states of this system. Property 5 γ (iv) guarantees subsequently that all energy differences between states of arbitrary systems have dyadic values, compared with the chosen unit of internal energy, and this implies that all systems are "U-measurable" with respect to the first system.

The relationships between Falk and Jungs metrization axioms and ours being clarified to some extent, a few critical remarks can be made.

Falk and Jungs metrization axioms are certainly too strong for the purpose for which they are meant. The authors themselves are conscious of the possibility of restricting certain properties to special systems, which will be used as energymeters, and that the boundedness of the domain of energy values requires another weakening of the axioms (notes p.132 and p.142).

The property (vi) is physically acceptable, only if we assume that non-dyadic energy values are indistinguishable from approximating dyadic energy values. If we wish to maintain the distinguishability, then the halving procedure does not define energy values of non-dyadic energy classes. In this case we are forced to have recourse to refined theories of the type developed in the third chapter.

Die Struktur der Thermodynamik : I Entropie und Energie

Axiom 1, and 2, and the additional supposition, made by Falk and Jung, that thermodynamic systems possess only one empirical entropy (p.136), are equivalent to the definition of adiabatic isolations as simple preorder relations on the set of systems $\mathcal{Z}$(axiom 1.6 in our approach). The metrization axioms for entropy are, according to Falk and Jung, contained in Axiom 3a, b, and the additional assumption that the empirical entropies of the systems $\mathfrak{z}_i$ are continuous variables. The latter assumption is equivalent to one of the properties of the entropymeter in our approach, namely $(\mathcal{S}, \rightarrow) \simeq (R', <)$. We may suspect that Axiom 3a, b and the last assumption are too strong for their purpose. Indeed, Axiom 3b cannot be reconciled with bounded additive entropy functions and the criticism formulated above applies again. A comparison between our metrization axioms for entropy and those of Falk and Jung, makes it apparent that axiom 3c is equivalent to the part of our axiom 2.9. i

$$(\forall z_i', z_i'' z_j)\ (z_i' z_j \rightarrow z_i'' z_j \text{ iff } z_i' \rightarrow z_i'')$$

The other part ... $z_i' z_j \leftrightarrow z_i'' z_j$ iff $z_i' \leftrightarrow z_i''$ belongs to those metrization axioms which are already implied in Axiom 3a, 3b and the continuity assumption with respect to the empirical entropy.

With respect to the other metrization axioms in our approach we again find it difficult to discover equivalent counterparts. Presumably remarks similar to those we made in the case of the energetic isolation hold here also.

In Axiom 4, which gives a final formulation of the properties of energetic isolations, the problem poses itself why part 4c is included. Axiom 4a and 4b are isomorphic with 3a and 3b, and because the axioms 2, 3a and 3b (and the continuity of the empirical entropy) imply, according to Falk and Jung, the metrization axioms for entropy, the axioms 2, 4a and 4b are expected to supply the metrization axioms of energy, and thus Axiom 4c. The crux is perhaps that an empirical energy is not defined and the continuity assumption for empirical entropy has no counterpart in the energy case. But it is difficult to believe that after 4a and 4b the complete set of properties 4c has to be introduced axiomatically.

Axiom 4b can again be rejected as physically irreconciliable with bounded additive energy functions.

Axiom 5 can be considered as partially superfluous, if extremal principles for entropy and energy are introduced, as Falk and Jung intend to do in Axiom 6.

II Die thermodynamischen Koordinaten

It is necessary to summarise the content of this part, because the argument is, in many places, ambiguous. It is, for instance, unclear whether conditions in implications are considered as sufficient or necessary.

A fair account of the main lines of this chapter is perhaps the following:

The authors start with a proof of the following theorem (not explicitly stated):

If (1) $X_1^{(1)} \ldots X_{n1}^{(1)}$ is a coordinate system of $\mathfrak{Z}_1$ and
$X_1^{(2)} \ldots X_{n2}^{(2)}$ is a coordinate system of $\mathfrak{Z}_2$,
$X_1^{(1)} \ldots\ldots\ldots\ldots X_{n2}^{(2)}$ being metric coordinates

and (2) $X_j^{(1)} + X_j^{(2)}$, $X_1^{(1)} \ldots, X_{j-1}^{(1)}, X_{j+1}^{(1)}, \ldots,$
$X_{n1}^{(1)}, X_1^{(2)} \ldots. X_{j-1}^{(2)}, X_{j+1}^{(2)} \ldots, X_{n2}^{(2)}$
is a coordinate system of $[\mathfrak{Z}_1, \mathfrak{Z}_2]_{\Gamma_j}$, i.e., the system which consists of the systems $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$ in contact equilibrium with respect to the contact relation (connection) associated with the coordinates of the same kind $X_j^{(1)}$ and $X_j^{(2)}$

and (3) $S_1(X_1^{(1)} \ldots X_{n1}^{(1)})$ and $S_2(X_1^{(2)} \ldots X_{n2}^{(2)})$ are differentiable with respect to the variables X_ℓ , all $\ell \varepsilon 1, \ldots n_1$, $1 \ldots n_2$.

then (a) $z_1 z_2 \varepsilon [\mathfrak{z}_1, \mathfrak{z}_2]_{\Gamma_j}$ iff $\dfrac{\partial S_1}{\partial X_j^{(1)}} = \dfrac{\partial S_2}{\partial X_j^{(2)}}$

(b) if $z_1 z_2 \varepsilon [\mathfrak{z}_1, \mathfrak{z}_2]_{\Gamma_j}$ then $d S = o$ if

$d(X_j^{(1)} + X_j^{(2)}) = o$ and $d X_k^{(1)} = o$, $k \neq j$ and

$d X_\ell^{(2)} = o$, $\ell \neq j$.

A similar theorem is generated by replacing S_1, S_2 and S by U_1, U_2 and U.

Axiom 6 states next that:

for all thermodynamic systems the condition (1) can be fulfilled (a), the condition (2) is fulfilled (b) and <u>necessary</u> conditions for $z_1 z_2 \varepsilon [\mathfrak{z}_1, \mathfrak{z}_2]_{\Gamma_j}$ are

(4) $S(z_1 z_2)$ = the maximum of $S(X_j^{(1)})$ under the conditions $X_j^{(1)} + X_j^{(2)}$, $X_k^{(1)}$ (all $k \neq j$), $X_\ell^{(2)}$ (all $\ell \neq j$) constant

(5) $U(z_1 z_2)$ = the minimum of $U(X_j^{(1)})$ under the conditions $X_j^{(1)} + X_j^{(2)}$, $X_k^{(1)}$ (all $k \neq j$), $X_\ell^{(2)}$ (all $\ell \neq j$) constant (b)

and finally adds (to make the theory consistent) that

(6) The maximum $\bar{S}(X_j^{(1)})$ and the minimum $\bar{U}(X_j^{(1)})$ will be reached at the same $X_j^{(1)}$ values (b)

(6*) The latter statement is changed in the case of systems whose energies have an upper and a lower bound: the maximum $\bar{S}(X_j^{(1)})$ and the extrema $\bar{U}(X_j^{(1)})$ will be reached at the same $X_j^{(1)}$ values (Ergänzung).

Finally, a thermal system (thermisches System) is defined as a system with a coordinate system U, $X_2 \ldots X_n$ and for this kind of system it is proved that:

Satz 10 : $$\frac{\partial S(U, X_2, \ldots X_n)}{\partial U} \equiv {}^{1}\!/_{T} > o$$

Satz 11 : Also S, $X_2 \ldots X_n$ is a thermodynamic coordinate system.

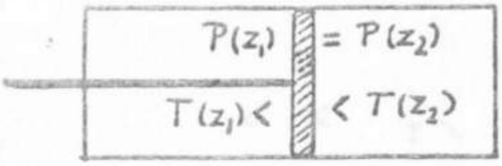
The criticism of this chapter is twofold:

The proofs of the theorems which are summarised above appear not to be sound : the conclusions cannot be proved on the basis of the given assumptions.

The content of axiom 6 has to be rejected on physical grounds insofar as the parts (4), (6) and (6*) are concerned.
In the case of assumption (4) it will be sufficient to give a counter example with a physical system for which assumption (4) clearly does not hold.

Counter example:

Consider a rigid container consisting of two compartments separated by a movable adiathermal piston. The compartments are filled with two quantities of an ideal gas.



Suppose that in an initial state

$z_1 \; z_2$:

$P\,(z_1) = P(z_2)$

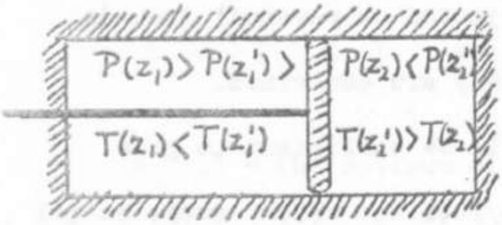
$T\,(z_1) < T\,(z_2)$

We will use the following notation: $U_{12}\,(z_1\; z_2) = U_1(z_1) + U_2(z_2)$

$S_{12}\,(z_1\; z_2) = S_1(z_1) + S_2(z_2)$ etc.

Now we move the piston quasistatically and under adiabatic isolation of the container until a state $z_1{}' \; z_2{}'$ is reached such that

$V(z_1{}') > V(z_1)$



In the state $z_1{}' \; z_2{}'$

$U_1(z_1{}') = U_1(z_1) + \Delta U_1$

$\Delta U_1 < 0$, thus $T(z_1{}') < T(z_1)$

because the gas is ideal

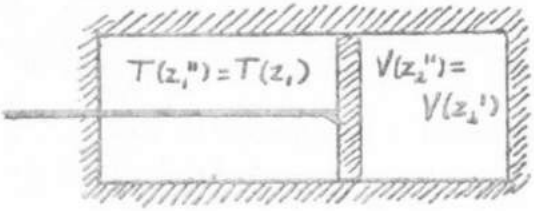
$U_2(z_2{}') = U_2(z_2) + \Delta U_2$

$\Delta U_2 > 0$, $T(z_2{}') > T(z_2)$

$\Delta U_{12} = \Delta U_1 + \Delta U_2 > 0$

$\Delta S_{12} = 0$, $\Delta S_1 = 0$, $\Delta S_2 = 0$

Then we establish thermal contact between the two compartments until a state $z_1{}'' \; z_2{}''$ is reached such that $T(z_1'') = T(z_1)$



$V(z_1{}'') = V(z_1{}')$

$V(z_2{}'') = V(z_2{}')$

$U_1(z_1'') = U_1(z_1') + \Delta U_1' = U_1(z_1)$, thus $\Delta U_1' = - \Delta U_1 > o$

and $T(z_1'') = T(z_1)$ because the gas is ideal

$U_2(z_2'') = U_2(z_2') + \Delta U_2'$ $\Delta U_2' = - \Delta U_1' = \Delta U_1 < o$

and $T(z_2) < T(z_2'') < T(z_2')$

Then:

$$\Delta S_1' \equiv S_1(z_1'') - S_1(z_1') > \frac{\Delta U_1'}{T(z_1)} = \frac{- \Delta U_1}{T(z_1)}$$

$$\Delta S_2' \equiv S_2(z_2'') - S_2(z_2') > \frac{\Delta U_2'}{T(z_2'')} = \frac{\Delta U_1}{T(z_2)}$$

$$\Delta S_{12}' > \frac{- \Delta U_1}{T(z_1)} + \frac{\Delta U_1}{T(z_2)}$$

Finally we establish thermal contact between compartment 2 and the environment until a state z_1'' z_2''' is reached such that

$U_2(z_2''') = U_2(z_2'') + \Delta U_2'' = U_2(z_2)$ thus

$\Delta U_2'' = - \Delta U_{12} = - \Delta U_1 - \Delta U_2 < o$

Thus $T(z_2''') = T(z_2)$

$$\Delta S_2'' \equiv S_2(z_2''') - S_2(z_2'') > \frac{\Delta U_2''}{T(z_2''')} = \frac{- \Delta U_1}{T(z_2)} + \frac{- \Delta U_2}{T(z_2)}$$

$(\Delta S_1'' = o)$

Now we compare the states z_1 z_2 and z_1'' z_2'''

$U(z_1'') = U(z_1)$ $\qquad$ $U(z_2''') = U(z_i)$

$V(z_1'') > V(z_1)$ $\qquad$ $V(z_1'') + V(z_2''') = V(z_1) + V(z_2)$

$$S_{12}(z_1'' \, z_2''') = S_{12}(z_1 \, z_2) + \Delta S_{12}' + \Delta S_2'' >$$

$$S_{12}(z_1 \, z_2) + \frac{- \Delta U_1}{T(z_1)} + \frac{- \Delta U_2}{T(z_2)}$$

or $S_{12}(z_1'' \, z_2''') - S_{12}(z_1 \, z_2) > -\frac{\Delta U_1}{T(z_1)} - \frac{\Delta U_2}{T(z_2)} =$

$$- \Delta U_1 \left(\frac{1}{T(z_1)} + \frac{\Delta U_2}{\Delta U_1} \frac{1}{T(z_2)} \right)$$

For the initial state, $P(z_1) = P(z_2)$.

For small changes dV, $dU_1 = P_1 \, d \, V_1$, $dU_2 = P_2 \, d \, V_2$,

thus starting from $z_1 \, z_2$ $dU_1 = P \, d \, V_1$ $dU_2 = -P \, d \, V_1$, and

$$\frac{dU_2}{dU_1} \simeq -1$$

Thus for small $d \, V$ starting from $z_1 \, z_2$:

$$d \, S = - dU_1 \left(\frac{1}{T(z_1)} - \frac{1}{T(z_2)} \right)$$

or : if $dU_1 < o$ and $T(z_2) > T(z_1)$ then $d \, S > o$

This result contradicts the assumption of Falk and Jung that the state $z_1 \, z_2$ is a state of maximal entropy compared with other states $z_1'' \, z_2'''$ for which $U(z_1'') = U(z_1)$, $U(z_2''') = U(z_2)$ and $V(z_1'' \, z_2''') = V(z_1 \, z_2)$.

The maximum entropy principle can be used only for states $z_1 \, z_2$ which are also in thermal equilibrium. This is clearly neglected by Falk and Jung.

If we reject assumption (4) then assumption (6) becomes unnecessary and meaningless.

Assumption (6*) has also to be rejected : Falk and Jung confuse upper and lower bounds for the energy of particular systems $_i$, with conditional maxima and minima for compound systems $[\mathfrak{Z}_i, \mathfrak{Z}_j]$

Finally, this criticism removes the justification of a more general theory of thermodynamics with special applications to thermal systems.

CHAPTER 8: A CRITICISM OF R. GILES, MATHEMATICAL FOUNDATIONS OF THERMODYNAMICS, Pergamon, 1964.

Starting point of the discussion of Giles' axiomatisation will be his formal theory, as explained in Appendix A p.191-214. This formal theory is independent of its physical interpretation. I will, however, add some remarks on the interpretation of the primitive terms and the axioms of the theory, for this will be necessary if we are interested in the question whether the formalism is an axiomatisation of traditional thermodynamics. Thus also, the foregoing chapters of Giles' work, where interpretations and justifications are given, will enter this discussion. Further, I will undertake a comparison with my own approach which is called C P T (axiomatisation of Classical Phenomenological Thermodynamics) and C P T O (the extension to open systems).

A reference to foregoing chapters of this thesis is given by C P T followed by the number of the chapter, definition, axiom or theorem.

"States": a, b, c .....

Giles' axiomatisation is an attempt to formulate a theory of the greatest possible generality. The starting point is a set $\mathcal{S}$ of states, which may be interpreted as physical states, comprising non equilibrium states. The 'rule of interpretation' has a certain vagueness:

"the state of a system represents its method of preparation ...., ... two states .... need not be distinguished if they are equivalent in respect of any prediction which might be made ......, ... the term state can refer only to conditions in which the system concerned is isolated." (1.4. p.17).

The second quotation reminds us that in physics the concept "variable of state" is logically prior to "state". It is however admissible to convert the logical order in an axiomatisation.

"Processes" α, β, γ, ...

The general concept 'process', being an ordered pair of states, has no physical interpretation. Only certain specified processes, e.g. 'natural processes', can be interpreted.

"Addition of states: +" and "Natural processes: →"

The interpretation is dependent on the axioms, which explain the use of the operators.

$a \rightarrow b$

A rule of interpretation for natural processes is given (1.5. p.24): " ... We write $a \rightarrow b$ if there exists a state k and a time interval t such that $a + k$ evolves (in isolation) in the time t into the state $b + k$".

The meaning of 'isolation' becomes clear in the definition of 'component of content' (A.3.2) and the enumeration of the quantities which appear as component of content, and remain constant in natural processes: the energy, the quantities of the chemical elements, the electric charge and the magnetic flux (if neutralisation of charge or current is not allowed), and volume. The equivalent of "natural process" in traditional thermodynamics is "process realisable under complete isolation and constant volume." This implies that the initial states of natural processes are either non equilibrium states, or states of systems which contain partitions and consequently they are 'possible non equilibrium states'.

$a + b \rightarrow c + d$

The interpretation is: 'natural process, where a evolves into c and b evolves into d'; Giles proceeds: "thus the process (a, c) can drive the process (d, b) backwards." (2.3. p.33).

One is tempted to interpret the processes (a, c) and (d, b) as physically meaningful. But the commutativity of a + b (Axiom A.2.1 (i)) implies that in this case also $b + a \to c + d$, thus (b, c) can drive (d, a) backwards. Processes, which are able to drive backwards other processes can easily be 'impossible processes', e.g. duplets of states of quite different systems. A too physical interpretation is thus not allowed.

The differences between Giles' axiomatisation and C P T do not exclude a comparison. The structure $(\mathcal{S}, +, \to)$ has similarities with the structure $(Z_i, +, \Rightarrow)$, where $z_i' + z_i''$ is identified with $z_i' z_i''$. The adiabatic accessibility relation $\Rightarrow$ is however quite different from $\to$ in its interpretation. The similarity will be clear if we compare the axioms A. 2.1, and A. 2.2 with the axioms in C P T.

A.2.1 Axiom.

(i) The operation + is associative and commutative

(ii) $a \to a$

(iii) if $a \to b$ and $b \to c$ then $a \to c$

(iv) $a \to b$ iff $a + c \to b + c$

A.2.2 Axiom

If $a \to b$ and $a \to c$ then $b \to c$ or $c \to b$

C P T Axiom 1.6, Axiom 2.9

For all $Z_i \in \mathcal{Z}$ there exist simple preorder relations $\Rightarrow_i$ such that:

(i) $z_i' \Rightarrow_i z_i''$ iff $z_i' z_j \Rightarrow_{ij} z_i'' z_j$

(ii) $z_i^1 \ldots z_i^n \Rightarrow$ Permutation $z_i^1 \ldots z_i^n$

(iii) if $z_i' \ldots z_i' \Rightarrow z_i'' \ldots z_i''$ then $z_i' \Rightarrow z_i''$

The only formal difference is that in C P T commutativity is restricted to states of the same system, and that C P T 2.9. (iii) has no equivalent.

"Components of content": $Q(x)$

Its physical interpretation as the internal energy, the deformation coordinates, and the elementary quantities has been given. Paragraph A.3 will demonstrate the existence of a positive component of content $Q(x)$ defined on all $x \in \mathcal{S}$, and will prove that if (a, b) is not possible, then there exists a positive component of content, such that $Q(a) \neq Q(b)$. The proof applies the extension theorem (B.3.1. p.219). An additive function $Q' (t\,\delta + \gamma) = t\,\lambda(\delta)$ defined on a subgroup $\mathcal{P}_\delta$ of processes $t\,\delta + \gamma$, where t is an integer, and γ is an arbitrary possible process, can be extended to an additive function Q'', defined on all processes, because $\lambda(\delta)$ is so defined that $Q' (t\,\delta + \gamma) \leq \lambda\,(t\,\delta + \gamma)$ and $\lambda(\alpha)$ is finite linear on the group of all processes $\mathcal{G}$, and further $\mathcal{P}_\delta$ contains an internal point ε with respect to λ. Next $Q(x) \equiv Q''\,(x, 2x)$ can be proved to be a positive component of content, and $Q(a) \neq Q(b)$ if (a, b) is not possible.

The definition of the function λ requires the introduction of an inclusion relation $\subset$, through:

A.3.3 $a \subset b$ iff there exists a positive integer n and a state c, such that $(n\,a + b, n\,b)$ is a possible process.

The existence of λ requires an axiom:

A.3.6 Axiom. There exists an internal state
A.3.5 Definition. A state e is internal if, given any state x, there exists a positive integer n such that $x \subset n\,e$.

The physical interpretation of the concept of an internal state causes some hesitation: it must be a state of a system containing an amount of all possible components of content: i.e. all chemical elements, volume, electric charge, magnetic flux, internal energy. The assumption of the existence of such a state might be avoided, and what kind of existence is meant?

Another criticism of the procedure outlined above, links up with a remark of Duistermaat ((6) p.21): "Let $\mathcal{R}$ be a subset of a commutative group $(G, +)$ and let $P_{\mathcal{R}}$ be the sub-semigroup of $(G, +)$ generated by $\mathcal{R}$ and 0. Then the largest subgroup of $(G, +)$ on which additive extensions of additive functions on $P_{\mathcal{R}}$ are uniquely determined is the group $G_{\mathcal{R}}$ of all $a \in G$ with $na = b-c$ for some natural number n and some $b, c \in P_{\mathcal{R}}$. Because of the arbitrariness of extensions of real measures of $\mathcal{R}$-processes to elements which do not belong to $G_{\mathcal{R}}$ ... it would be reasonable to extend measures of $\mathcal{R}$ processes only to $G_{\mathcal{R}}$ and not to the whole group G." This implies that the largest subgroup of $\mathcal{P}$ on which an additive extension Q'' of additive functions Q' on $\mathcal{P}_\delta$ is uniquely determined is the group $\{\alpha: n\,\alpha = t\,\delta + \gamma\}$. If a state a is chosen as unit of content: $Q(a) = Q''(a, 2a) = Q''(\delta) = 1$, then $Q(x) = Q''(x, 2x)$ is determined only if $n(x, 2x) = t \cdot (a, 2a) + (p,q)$ where (p,q) is possible, or $n\,x = t\,a + p$ and $2\,n\,x = 2\,t\,a + q$. If we represent x, a, p and q by vectors in content space $\underset{\sim}{Q}(x)$, $\underset{\sim}{Q}(a)$, $\underset{\sim}{Q}(p)$ and $\underset{\sim}{Q}(q)$, then, because $\underset{\sim}{Q}(p) = \underset{\sim}{Q}(q)$, it follows $n\,\underset{\sim}{Q}(x) = t(\underset{\sim}{Q})(a)$ which implies that x and a differ only in a factor $t/_n$ in all components of content, thus x and a are 'similar states' in the terminology of C P T O. The domain of definition is thus indeed utterly restricted, and the mathematical argument does not lead to a procedure which enables us to measure components of content, and the components of content defined through $Q(x) = Q''(x, 2x)$ do not seem physically relevant.

The proof that $Q(a) \neq Q(b)$ if (a,b) is not possible necessitates the introduction of another axiom:

A.4.3 Axiom. Given a process α, if there exists a state c such that for any positive real number ε there exist positive integers m, n and states x, y such that $^m/_n < \varepsilon$, $x \subset m\,c$, $y \subset m\,c$, and $(x,y) + n\,\alpha \to o$ then $\alpha \to o$.

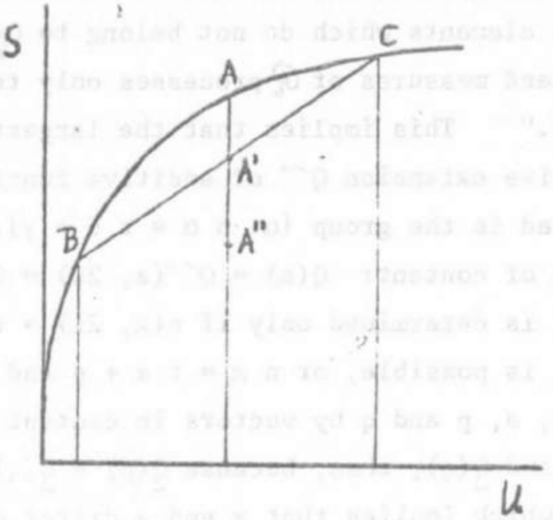


figure 8.1

The direct physical interpretation of the axiom is thwarted by the interpretation of $\subset$ and the expression $(x,y) + n\,\alpha \to o$ in case (x,y) and α are impossible processes. But something can be said in view of certain implications of the axiom.
Firstly: the Archimedean axiom is implied:

A.4.4 Theorem. Given states a, b if there exist states x,y such that the relation $n\,a + x \to n\,b + y$ holds for arbitrarily large integers n then $a \to b$.

This axiom enables us to include in the natural (or possible) processes those which are not natural (or possible) in themselves but can be 'driven' with the assistance of 'infinitesimal' processes. In C P T the <u>interpretation</u> of 'possible process' is so wide that these processes are included already.

Another implication of A.4.3 is, that for states a, b of the same 'content' the entropy difference $\Delta\, S(a,b)$ is finite (see A.4.5). Within the framework of C P T this leads to the interesting conclusion that the entropy S has a lower bound (an axiom in C P T, which is part of the third law: C P T 6.4). Consider the $S(U)_{X_k}$ curve of a thermodynamic system which obeys the maximum entropy principle, and continuity assumptions, and whose internal energy function has a lower bound (see figure 8.1). The points A, B, C on the curve represent equilibrium states a, b, c; the points A', A'', ... represent non-equilibrium states a', a'', ... with the same content as a, and consisting of states 'similar' to b,c (see C P T O (2) p), so that $a' = \frac{A'C}{BC}\ b + \frac{A'B}{BC}\ c$. A necessary condition for finiteness of $\Delta\, S(a, a')$ at constant content is finiteness of the distance $A\,A'$, and this implies in case of a lower bound of U, a lower bound for S.

Quasi-Entropy: $S(x)$

The concept of 'quasi-entropy' replaces the role of the concept of 'entropy' in a theory that gives 'natural processes' a more fundamental meaning than 'adiabatic processes'. In such a theory the connection with traditional thermodynamics will be established through additional assumptions concerning the 'quasi-entropy' of 'mechanical states' or the 'quasi-entropy difference' of 'mechanical processes'. The formal similarities of the quasi-entropy as defined in A.4.1 and the set of extensive entropy functions S_i as defined in C P T 2.1 will be clear after inspection.

A.4.1 Definition. A real valued function $S(a)$ defined for every state a is a quasi-entropy function if:

(i) $S(a + b) = S(a) + S(b)$

(ii) if $a \rightarrow b$ & $b \rightarrow a$ then $S(a) = S(b)$

(iii) if $a \rightarrow b$ & $b \not\rightarrow a$ then $S(a) < S(b)$

C.P.T. 2.1 Definition.

A set of extensive entropy functions S_i for the systems $Z_i \varepsilon \mathscr{Z}$ is a set of real valued functions $S_i(z_i)$, $z_i \varepsilon Z_i$, such that

(i) $S_i(z_i') \lesssim S_i(z_i'')$ iff $z_i' \Rightarrow z_i''$

(ii) if $Z = Z_i \times Z_j \times \dots$ then $S(\overline{\overline{z}}) = S(z_i\ z_j\ \dots) = S_i(z_i) + S_j(z_j) + \dots\dots$

The proof of the existence of a positive quasi entropy function $S(a)$ for all states $a \varepsilon \mathscr{S}$ is similar to the existence proof of a positive component of content and similar criticism applies. Again a special function, $\nu(\alpha)$, ii introduced. It is not profitable to consider its physical interpretation. With the help of this function a function $I(\alpha)$ is defined on the set of possible processes

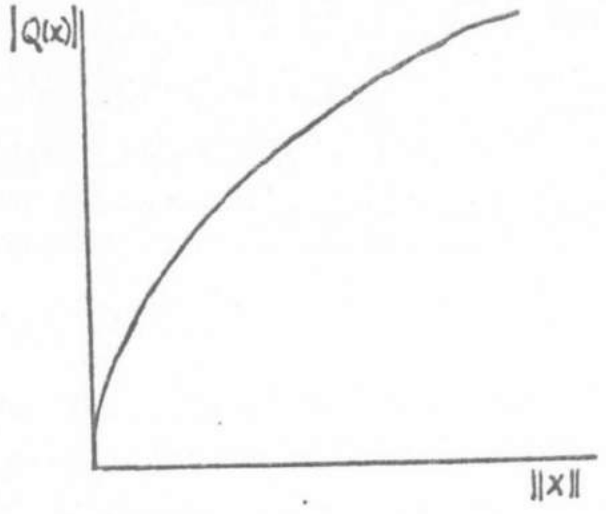


figure 8.2

$\mathcal{P}_p$ and this function can be extended to a function I′ defined for all processes. Finally, a function $S(a) \equiv I'(a, 2a)$ is defined for all states and it is demonstrated that S(a) is a positive quasi entropy function. The function I may be <u>formally</u> extended to all processes, it is, however, uniquely defined for the set of processes $\{\alpha : n\,\alpha = m\,\beta\}$, where β is a possible process ($\beta \in \mathcal{P}_p$). In C P T the introduction of functions, whose meaning seems to be exclusively mathematical, is avoided. An attempt to define internal energy functions and entropy functions on the domain of <u>all</u> states is not undertaken, because it seems unprofitable.

<u>Boundedness</u>

A.6.3 Definition. A real valued additive function Q(a) defined for every state a is bounded if there exists a constant k such that, for all a, $|Q(a)| \leq k\,||a||$.
($||a|| = \inf\{{}^m/n : na \subset me\}$, where e is the unit state (A.6.1))

The physical interpretation leads to difficulties owing to the use of the concepts 'internal state e' and 'inclusion $\subset$'. The justification of its introduction as given on p.63 is not convincing:

" ... a component of content Q is physically acceptable if and only if it is bounded .... If Q is not bounded then we can find states x with $||x||$ arbitrarily small but with $|Q(x)| > 1$, say. But then, for any state a, the states a and a + x differ arbitrarily little in content, and yet $|Q(a + x) - Q(a)| > 1$. On the other hand, if Q is bounded this cannot happen....." The following example demonstrates that boundedness is not necessary to avoid the situation described. Suppose the graphic representation of the function $||x|| \rightarrow |Q(x)|$ is as in figure 8.2, with $\lim_{||x|| \rightarrow 0} |Q(x)| / ||x|| = \infty$: this function is not bounded, but remains physically acceptable.

The requirement of 'boundedness' seems to be introduced to guarantee continuity of thermodynamic functions. Its counterpart in C P T might be the continuity assumptions of chapter 4.

Equilibrium states

A.7.1 Definition. A state a is an equilibrium state if there exists no state b such that $a \rightarrow b$ and $b \nrightarrow a$.

The definition implies that state a is a state of maximum entropy, compared with all other states with the same 'content'. This clarifies the relation of Giles' axiomatization and Gibbsian thermodynamics, and consequently with C P T, where 'equilibrium state' is a primitive concept and the axiom can be stated that equilibrium states are states of maximum entropy. Because in C P T a more specified concept of equilibrium is given through the distinction of different 'connections', the maximum entropy principle can be stated also in a more specified form (see C P T 5.1).

A.7.2 Definition. A state is a perfect equilibrium state if na is an equilibrium state for every positive integer n.

Giles relates this concept in its physical interpretation with equilibrium states of simple (homogeneous) fluids. The associated concept in C P T O is 'equilibrium state of a homogeneous system'. In C P T O 'homogeneity' is firstly defined and next the theorem is proved: for all 'similar states' z_i and z_j: $\frac{S_i(z_i)}{S_j(z_j)} = \frac{U_i(z_i)}{U_j(z_j)} = \frac{V_i(z_i)}{V_j(z_j)} = \frac{M_i}{M_j}$ ((2) p.   )

The converse of this theorem is : for homogeneous sytems states a and r a are 'similar', and consequently if a is an equilibrium state, then r a is also.

Potential :

A.8.1. Definition. A potential $\underset{\sim}{\varphi}$ is an additive function on content space $\mathcal{Q}$ such that $-\underset{\sim}{\varphi}(\underset{\sim}{Q}(x)) \geq S(x)$ for every state x.

A.8.2. Definition. We say state a has the potential $\underset{\sim}{\varphi}$ if $S(a) = -\underset{\sim}{\varphi}(\underset{\sim}{Q}(a))$.

It is not allowed to interpret a potential as a vector consisting of absolute temperature and Massieu functions. This interpretation would pre-suppose differentiability of the function $S(\underset{\sim}{Q}(x))$. Giles explains what can be derived without the assumption of continuity, and he realises that more explicit results can be obtained if such assumptions are used. The final result of paragraph A.8 is:

A.8.5. Theorem. If c = a + b is a perfect equilibrium state having the potential $\underset{\sim}{\varphi}$ then a and b are themselves perfect equilibrium states having the potential $\underset{\sim}{\varphi}$, and conversely.

In C P T assumptions of differentiability are indeed introduced and these lead to the existence of absolute temperature and Massieu functions and to the theorems C P T 5, 11-17.

$$z_i \, z_j \in C_\theta \text{ iff } T(z_i) = T(z_j)$$

$$z_i \, z_j \in C_\theta \text{ and } z_i \, z_j \in C_{\phi k} \text{ iff } F_{k/T}(z_i) = F_{k/T}(z_j)$$

which, together with the transitivity of the connection relations may be considered as specifications of the above theorem A.8.5.

Absolute entropy

The last paragraph of Giles'axiomatization introduces the concept 'anti-equilibrium state' and assumes the existence of an anti-equilibrium state for every state x, with the intention to obtain in this way states of reference for an absolute entropy function:

A 9.1. Definition. A state x is an anti-equilibrium state if there exists no state a such that $a \to x$ and $x \not\to a$

A 9.4. Axiom.

(1) If a is any state then there exists an anti-equilibrium state such that $x \to a$

(2) If x and y are anti-equilibrium states then so is $x + y$

A 9.3 Definition. A function of state S is an absolute entropy function if:

(1) S is a quasi-entropy function

(2) $S(x) = o$ for every anti-equilibrium state x

The axiom A 9.4 is physically very dubious. It could be argued that the contrary is physically more realistic: such an axiom would express that states of zero entropy and zero absolute temperature do not exist; the points in phase space, which represent such states, form the closure of the <u>open</u> set of 'occupied' points.

Giles' approach is essentially different from that of C P T. In C P T no attempt is made to define an absolute entropy function on the set of all states. For certain subsets of states (formal systems) a greatest lower bound for the entropy is assumed and it is recognised as part of the 'third law' of traditional thermodynamics.

Apart from a 'quasi-entropy' and an 'absolute entropy', Giles introduces an 'entropy function':

A 9.12. Definition. A function of state S_o is an entropy function if

(i) S_o is a quasi-entropy function, and

(ii) $S_o(m) = o$ for every mechanical state m.

The mechanical states form an arbitrary subset of the anti-equilibrium states. A rule of interpretation is not given and Giles is aware of the difficulties involved. Duistermaat has rightly remarked that it will be physically more acceptable if we omit 'mechanical states' and restrict ourselves to the introduction of 'mechanical processes'((6) p.51): this also avoids the problem of the connection between 'mechanical states' and 'anti-equilibrium states'.

Construction of an energy function

Considering the failure of the theory, as described above, to construct component of content functions which are uniquely defined on physically interesting subsets of $\mathcal{S}$, it is desirable to discuss the construction of an energy function which Giles offers in paragraph 11.3.

The procedure is as follows:

(1) Select two arbitrary states A_o and A_1 of a system A, and <u>define</u> $E(A_o) = o$, $E(A_1) = 1$

(2) Prepare a row of systems in state A_o and a row in state A_1. To determine the energy value of a state A_2, which is intermediate in hotness between A_o and A_1, weak thermal contact between A_o and A_1 is established, till one of the systems reaches the state A_2. As soon as that happens this system is replaced by a second system of the same row, etc.

After m + n steps : $m\,A_1 + n\,A_o \rightarrow (m + n - 1)\,A_2 + A_3$

Then, according to Giles, also A_3 is intermediate in hotness between A_o and A_1; thus $A_2 \subset A_o + A_1$, and $A_3 \subset A_o + A_1$; consequently $||A_2|| \leq ||A_o + A_1||$, and $||A_3|| \leq ||A_o + A_1||$; thus

$$E(A_2) = \lim_{m + n \rightarrow \infty} \left(\frac{m}{m + n} \quad E(A_1) \quad + \quad \frac{n}{m + n} \quad E(A_o) \right]$$

This procedure is only applicable to one dimensional systems (or two-dimensionality of the 'completion'), which means two systems with internal energy as completely determining variable of state. It is however not clear which criterion leads to the selection of the subsets of states forming these systems, from the set $\mathcal{S}$ of all states. It pre-supposes a more specified theory than the general theory of ($\mathcal{S}$, +, →).

Further, the maximum entropy principle (or an equivalent statement) is presupposed, and presumably also differentiability of $S(E)$. Finally, boundedness of E is presumed, which property seems not clear physically. This criticism will lead to an alternative approach in C P T, which presumes a division of $\mathcal{S}$ in subsets Z_i, or 'systems', and a selection of systems with special properties : energy meters (see C P T 3.15,16).

Construction of an entropy function

This chapter will close with a few remarks about Giles' construction of an entropy function in paragraph 11.4. This construction might be reproduced as follows:

(1) Select two arbitrary states A_o and A_1 of a system A, and define $S(A_o) = o$ and $S(A_1) = 1$.

(2) An auxiliary system K and a mechanical system M are necessary to determine the entropy of an arbitrary state A_2 of the system A. A series of quasistatic adiabatic processes of systems A + K, and quasistatic natural processes of K + M is performed, such that the result is:

$m\,A_o + n\,A_1 + K_o + M_1 \Longleftrightarrow (n + m)\,A_2 + K_1 + M_2$, and next the natural irreversible process $K_1 + M_2 \rightarrow K_o + M_3$.

It follows that $m\,A_o + n\,A_1 + M_1 \rightarrow (n + m)\,A_2 + M_3$, where (M_1, M_3) is a 'mechanical process'. Thus $S(A_2) \geq \frac{m}{m+n} S(A_o) + \frac{n}{m+n} S(A_1)$, and because only the last step is irreversible natural:

$$S(A_2) = \lim_{m+n \rightarrow \infty} \left[\frac{m}{m+n} S(A_o) + \frac{n}{m+n} S(A_1) \right] .$$

The procedure is again only applicable for one-dimensional systems. It assumes the concept of a quasistatic adiabatic process, which falls outside the framework of the general theory of $(S, +, \rightarrow)$. Giles argues that the necessity of quasistatic performance of the processes is superfluous: ' .... we may not be prepared that O (i.e. a 'primitive observer') can recognise equality of hotness. However, in principle this is not necessary. For if O carried out the experiment (for a given value of $N = n + m$) a large number of times, making the said adjustment at random, it will occasionally happen by chance that their accuracy is good; since the experiment determines a bound to the unknown quantity, rather than an estimate of it, one good trial supersedes any number of bad ones'. This argument is invalid, because in that case each of the N process steps of $A + K + M$ adds to the irreversibility and if $N \to \infty$, then the probability p to obtain a good total accuracy approaches zero.

In C P T there is a possibility to choose 'one dimensional systems'. Also the concept of a 'quasi-static adiabatic process' is available. The construction of an entropy function in C P T (2. 16, 17) is, in some respects, similar to Giles' construction, in other respects it is different: e.g. following Cooper (5) another limiting procedure is used.

CHAPTER 9: A CRITICISM OF J. J. DUISTERMAAT, ENERGY AND ENTROPY AS REAL MORPHISMS FOR ADDITION AND ORDER, Synthese 18, 311 (1968).

Duistermaat's article will be discussed in so far as thermodynamics is considered. Therefore, I will start from part III:

Thermodynamics. The given axiomatisation will be considered with respect to its consistency, and with respect to its physical interpretation. Also, a comparison will be made with C P T.

3.1.1. Axiom: $(S, +, \mathcal{N})$ is a preordered commutative semigroup.

$(S, +, \mathcal{N})$ has to be interpreted as $(\mathcal{S}, +, \rightarrow)$ in Giles' axiomatisation. The same criticism applies. The axiom expresses that $(S, +, \mathcal{N})$ satisfies the conditions:

1.2.6 (ii) $(a + x, b + x) \in \mathcal{N}$ iff $(a, b) \in \mathcal{N}$

1.2.6 (iii) $\mathcal{N}$ is a preorder in S (i.e. reflexivity and transitivity applies).

This axiom is thus equivalent with Giles' axiom A.2.1., which is already compared with C P T. (see chapter 8)

3.1.2. Axiom: There exist irreversible $\mathcal{N}_{as}$-processes.

The concept of an $\mathcal{N}_{as}$-process is explained in 1.3.5. With the commutative semigroup of states $(S, +)$ is associated the commutative group of processes $(S \times S, +)$, where addition is defined by $(a, b) + (c, d) = (a + c, b + d)$. $S \times S$ falls apart in equivalence classes $[a, b]$ through the equivalence relation $\sim$, defined as follows (1.1. p.10):

$(a, b) \sim (a', b')$ iff $(\exists\, x, x')\ (x, x' \in S\ \&\ (a + x, b + x) = (a' + x', b' + x'))$. The introduction of the group $S \times S$, and the equivalence relation $\sim$ seems to be performed because of its mathematical usefulness, it can be doubted whether a physically meaningful interpretation for $\sim$ is possible (see C P T chapter 8).

Let us, for clarity, indicate states by roman letters a, b, ... and formal processes by greek letters α, β, ... Interest is restricted to certain subgroups of the group of formal processes $\mathcal{G} = S \times S$, namely to the natural processes $\mathcal{N}$, the mechanical processes $\mathcal{M}$ the adiabatic processes $\mathcal{A}$. An attempt to formulate a general theory for the group of all formal processes $\mathcal{G}$ seems to me unprofitable, as already explained in chapter 8. Entropy difference functions and energy difference functions can only be sufficiently defined on subgroups of $\mathcal{G}$.

In the group $(\mathcal{G}, +)$ a preorder $\lesssim$ is generated by a subgroup $\mathcal{N}$ of natural processes: $\alpha \lesssim \beta$ iff $\beta - \alpha \in \mathcal{N}$ (if $\alpha = (a, b)$ then $-\alpha \equiv (b, a)$). $\alpha \lesssim \beta$ can be interpreted as 'the process β can drive the process α backwards'. This has the usual meaning, if α and β are possible processes. This restriction is, however, not made by Duistermaat and the criticism of chapter 8 can be repeated.

The subgroup $\mathcal{N}$ is extended to a subgroup $\mathcal{N}_{as}$, which is introduced to comprise those processes which are not natural in the strict sense (i.e. spontaneous under complete isolation) but which can be performed with the help of an 'arbitrarily small' spontaneous process.

The author explains that, if the existence of a 'dominating process' is assumed (i.e. a process α such that for all $\xi \gtrsim 0$, which implies $\xi \in \mathcal{N}$, there exists a natural number n such that $\xi < n\,\alpha$), $\xi \underset{as}{\lesssim} y$, iff $(\forall r)\ (\exists n, k)\ ({}^{n}/k < r \ \&\ k\,(\xi - y) \lesssim n\,\alpha)$, where r is a positive real number, and n and k are natural numbers. The assumption of a 'dominating process' is closely related to the assumption in C P T that all adiabatic processes are measurable by a unit process of an entropy meter and similar assumptions for processes under energetic isolation (see C P T 2.21-22, 3.18-20). The extension of $\mathcal{N}$ to $\mathcal{N}_{as}$ has a function which is parallel with the introduction of the Axiom A.4.3. in Giles' axiomatisation and, as is already mentioned in chapter 8, the problems solved on the formal level through it by Giles and Duistermaat are solved on the level of physical interpretation of $\Rightarrow$ and $\lesssim$ u) in C P T, where reversible accessible processes include processes which are not spontaneous but 'quasistatic' accessible under the given isolation. This corresponds with an approach in which $\lesssim$ and $\lesssim_{as}$ are assumed to be identical, i.e. the preordered commutative group $(\mathcal{G}, +, \lesssim)$ is "Archimedean" (see (6) 1.3.8). The axiom 3.1.2.: "there exist irreversible $\mathcal{N}_{as}$-processes" is associated with: "there exist states z_o', z_o'' of a meter Z_o such that $z_o' \rightarrow z_o''$" which is implied in C P T 2.17.

The axioms 3.1.1 and 3.1.2 imply, according to Duistermaat, the existence of non-zero measures of natural processes. This is presumably to be considered as an application of the corollary of theorem 1.4.1.

1.4.1. Theorem $(a, b) \in \mathcal{R}_{as}$ iff $f(a) \lesssim f(b)$ for each $(S, +, \mathcal{R})$ morphism f.

Corollary: There exists a non trivial real $(S, +, \mathcal{R})$ morphism iff there exists an irreversible $\mathcal{R}_{as}$-process.

But it can be doubted that the corollary follows from the theorem. A possible proof might be attempted in this way:

$(a, b) \in \mathcal{R}_{as}$ & $(b, a) \notin \mathcal{R}_{as}$ iff $f(a) \lesssim f(b)$ and not $f(b) \lesssim f(a)$ for each real $(S, +\mathcal{R})$ morphism f; this does not imply $(\exists f)\ (f(b) > f(a))$ because the set of real $(S, +, \mathcal{R})$ morphisms might be empty!

A complete treatment of the necessary and sufficient conditions for the existence of measures of $\mathcal{N}_{as}$-processes is given in the general theorem 1.3.4:

1.3.4. Theorem. A real additive function f_o on a subgroup G_o of a preordered commutative group $(G, +, \lesssim)$ can be extended to a real $(G, +, \lesssim)$-morphism if and only if f_o is dominated by some real function φ_o on some dominating subset A of G.

The foregoing conditions are fulfilled for $(\mathcal{G}, +, \mathcal{N})$ if for the set G_o is chosen the set of processes {n γ : n = integer, γ = unit process}, f (n γ) = n f(γ) = n, and a dominating function ν(α) is defined as in Giles' axiomatisation ((4) A.4.6). This approach has already been criticised in chapter 8. A simplification of this oan be obtained if we choose for G the subgroup of possible processes $\mathcal{G}_P$, for G_o again the set {n γ} and f(n γ) = n, and γ is considered as a dominating element. This is in fact the approach of C P T. It has not the generality of Giles' or Duistermaats' approach, but the latter recognises that extension of measures of $\mathcal{N}$ processes to the whole group $\mathcal{G}$ is not profitable (see 1.3.8).

The uniqueness of existing (G, +, $\lesssim$) morphisms on a subset $G' \subset G$ requires an extra condition (1.4.2): G' is linearly ordered with respect to $\lesssim_{as}$. This condition can be introduced as an axiom which, applied to $(\mathcal{G}, +, \mathcal{N})$, runs as follows:

3.1.3. Axiom. If $(x, y) \in \mathcal{N}_{as}$ and $(x, z) \in \mathcal{N}_{as}$ then $(y, z) \in \mathcal{N}_{as}$ or $(z, y) \in \mathcal{N}_{as}$

This axiom is closely related with Giles' axiom A.2.2. and with the assumed comparability and transitivity of the relation ⇒ in C P T. Duistermaat prefers however to derive 3.1.3 from the definition of an entropy function and two axioms, introducting the set of mechanical processes $\mathcal{M}$ as a new primitive term.

3.2.1. Axiom: (i) $\mathcal{M}$ is an equivalence relation in S

(ii) Catalysed sums of zero and $\mathcal{M}$ processes are $\mathcal{M}$ processes (i.e. $(a + x, b + x) \in \mathcal{M}$ iff $(a, b) \in \mathcal{M}$ and $\mathcal{M}$ is a preorder in S).

3.2.2 Definition: S is an entropy function on $(S, +, \mathcal{N}, \mathcal{M})$ iff:

(i) S is a real additive function on (S, +)
(ii) if $(a, b) \in \mathcal{N}$ then $S(a) \leq S(b)$
(iii) if $(a, b) \in \mathcal{M}$ then $S(a) = S(b)$

3.2.4 Axiom: Each mechanical process which is a difference of two natural processes is a reversible asymptotically natural process.

The derivation of 3.1.3 can be reproduced as follows_

3.2.4. , iff each measure of $\mathcal{N}$ processes can be extended to an entropy increase function Δ S on $\mathcal{G}$, iff measures of $\mathcal{N}$ processes are uniquely determined up to a gauge factor, iff 3.2.1.

The second step assumes that Δ S on $\mathcal{G}$ is uniquely determined up to a gauge factor. But it is not perfectly clear which presuppositions are involved in the proof of this statement as undertaken in 3.5. It might be preferable to take 3.1.3 above 3.2.4 as an axiom.

The discussion of the relationships between Duistermaat's axiomatisation and traditional thermodynamics is hindered by the lack of clear rules of interpretation for the primitives: the remark '"thermal isolation" seems to be an essential part of the physical interpretation of mechanical processes' (3.2. p.50) is insufficient. An attempt to give a rule of interpretation for $\mathcal{M}$ other than 'quasistatic adiabatic processes' leads into difficulties. It seems that the introduction of $\mathcal{M}$ with the properties of axioms 3.2.1 and 3.2.4 begs the question of the relationships between the physical content and formal structure of thermodynamics.

Energy. The First Law of Thermodynamics

The introduction of an internal energy function takes place in a manner essentially different from Giles' approach and can be summarised as follows:

Primitive terms are introduced: the set of purely mechanical natural processes: $\mathcal{N}_{\text{mech}}$, the set of all states of S which are 'at rest': S_o, the set of pairs of states which only differ with respect to their properties of motion: $\mathcal{M}_o$.
Next an energy function is defined:

2.2.2. Definition. E is an energy function on the mechanical system $(S, +, \mathcal{N}_{\text{mech}}, S_o, \mathcal{M}_o)$ iff;

(α) E is a real additive function on (S, +)

(β) If $(a, b) \varepsilon \mathcal{N}_{\text{mech}}$ then $E(a) = E(b)$

(γ) If $(a, b) \varepsilon \mathcal{M}_o$ and $b \varepsilon S_o$ then $E(a) \gtrless E(b)$

A passive process $\alpha \varepsilon \mathcal{P}$ is defined as a catalysed sum of $\mathcal{N}_{\text{mech}}$, stopping and zero processes, stopping processes being processes a, b $\varepsilon \mathcal{M}_o$ such that b ε S_o. Consequently E is an energy function iff − E is a real $(S, +, \mathcal{P})$- morphism.

Next a procedure for the experimental determination of E differences is explained (2.2.3). This procedure restricts itself to the measurement of energy changes op passive processes (which may be interpreted as processes in which the kinetic energy decreases, and no other energy changes occur). Then it is assumed (3.3 p.55) that the energy increase function Δ E is uniquely determined on $\mathcal{M}$ up to a gauge factor: this means that the domain of unique definition of the Δ E function is extended to processes for which, until now, Δ E could not be measured experimentally. A procedure for measurement in this domain is not given.

The last step is the definition of a thermodynamic energy function as an extension of an energy function from $\mathcal{N}_{\text{mech}}$ to $\mathcal{M}$ (3.3.1) which is made possible without inconsistencies by axiom 3.3.2.

3.3.1. **Definition:** E is called a thermodynamic energy function with respect to the mechanical energy increase function Δ E on iff:

(i) E is a real additive function on (S, +)

(ii) If $(a, b) \varepsilon \mathcal{N}$ then $E(a) = E(b)$

(iii) If $(a, b) \varepsilon \mathcal{M}$ then $E(b) - E(a) = \Delta E([a,b])$.

3.3.2. **Axiom:** If α is a mechanical process which is a difference of two natural processes, then $\Delta E(\alpha) = o$ for each mechanical increase function Δ E on $\mathcal{M}$.

Duistermaat considers this axiom as "the first law of thermodynamics". This seems incomplete: the assumed possibility to extend the mechanical energy increase function Δ E to mechanical processes is an equally important part of the first law, within the framework of this theory.

One parameter Contact Experiments. Heat. The Second Law of Thermodynamics

In this and the following paragraphs, the axiomatisation restricts itself to 'narcistic systems , parametrised by a component of content E'.

Definition: A subsystem T of S is called parametrised by E if $a, b \varepsilon T$, $E(a) = E(b)$ implies that $(a, b) \varepsilon \mathcal{N}_{as}$.

Definition: $T \subset S$ is called narcistic if for each $a, b \varepsilon T$ there exists a state $c \varepsilon T$ with $(a + b, c + c) \varepsilon \mathcal{N}_{as}$.

The only systems of traditional thermodynamics which obey the condition 'parametrized by E' non trivially, are systems which are completely determined by the value of the internal energy. These are systems without variable deformation coordinates or systems with deformation coordinates whose values are completely defined by the internal energy, e.g. a gas with variable volume but with constant temperature or constant entropy. This is a severe restriction on the generality of the theory, and it might be considered as inacceptable. For narcistic systems, parametrized by E, it is proved that S(E) is dyadically concave, but not necessarily continuous. Duistermaat assumes in subsequent sections that S(E) is continuous but not necessarily differentiable.

A concept of equilibrium is defined (p.60):

' ... T_1 and T_2 are in equilibrium at $a_1 \varepsilon T_1$, $a_2 \varepsilon T_2$, iff for each $\varepsilon > o$ there are $a_1' \varepsilon T_1$, $a_2' \varepsilon T_2$ with $E(a_1) + \varepsilon > E(a_1') > E(a_1)$ such that $(a_1' + a_2', a_1 + a_2) \varepsilon \mathcal{N}_{as}$, and there are $a_1'' \varepsilon T_1$, $a_2'' \varepsilon T_2$ with $E(a_1) > E(a_1'') > E(a_1) - \varepsilon$ such that $(a_1'' + a_2'', a_1 + a_2) \varepsilon \mathcal{N}_{as}$.'

This can be expressed as: $S(a_1 + a_2)$ is maximal compared with $S(a_1' + a_2')$, where a_1', a_2' are neighbouring states, such that $E(a_1' + a_2') = E(a_1 + a_2)$. This is clearly the maximum entropy principle, restricted to systems 'parametrized by E'. Narcistic systems can be interpreted as systems whose E space is 'dyadically occupied' and for which the maximum entropy principle holds good.

In C P T primitive terms 'thermal connection' or 'in thermal equilibrium', and 'k^th force connection', or 'in k^th force equilibrium' make it possible to replace the above definition of 'equilibrium' by a similar but more specified statement which has the status of an axiom. This specification, and the extension of the principle to 'simple systems' which are systems with multi dimensional phase space $[U, \dots X_k, \dots]$, are essential for the further development of a general theory of thermodynamics.

The subsequent theorem (3.4.4) reads: For subsystems T_1, T_2 of S, such that T_1 and T_2 are parametrized by E, and S(E) is differentiable:

T_1 and T_2 are in equilibrium at $a_1 \varepsilon T_1$, $a_2 \varepsilon T_2$ iff $\dfrac{\partial S(a_1)}{\partial E} = \dfrac{\partial S(a_2)}{\partial E}$.

This theorem is similar with C P T 5.10: under certain conditions ....

$$z_i\ z_j\ \varepsilon\ C_\theta \quad \text{iff} \quad \left(\frac{\partial S_i(z_i)}{\partial U_i}\right)_{X_{ki}} = \left(\frac{\partial S_j(z_j)}{\partial U_j}\right)_{X_{\ell j}} .$$

Next a new primitive term is introduced by Duistermaat: 'thermal system'. The restriction of an energy function to a thermal system T is called a 'heat function' on T. Thermal systems could be interpreted as systems with constant deformation coordinates, but the interpretation is not really attempted.

The second law is then announced in the following form:

Axiom: Let E be a thermodynamic energy function as in 3.4.5. Then $[a, a'']$ is an irreversible adiabatic process if a, a'' belong to the same thermal system $T \in \theta$ and if $E(a) < E(a'')$

Criticism of this axiom is possible in so far as it does not allow for negative absolute temperatures. The axiom holds only for 'thermal systems' with positive absolute temperature, and if we restrict ourselves to the latter class of systems then the axiom is an implication of the concave upwardness of the S(E) curve, and is consequently no axiom at all. It cannot be expected that a 'second law' would still be necessary after the foregoing.

The measurement of entropy differences as explained in 3.5. is done with the help of Carnot cycles. It presupposes at least the existence of such cycles adapted to arbitrary processes. According to Duistermaat all known measurements of entropy seem to be of this kind. This can be doubted. Alternative procedures are given by Giles (4) and Cooper (5), and followed in C P T.

CHAPTER 10: A CRITICISM OF J. L. B. COOPER, THE FOUNDATIONS OF THERMODYNAMICS. JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 17, 172-193 (1967).

The article consists of an introductory part with a critique of classical theories and a systematic development of an axiomatic system for thermodynamics.

The critique of classical theories is certainly the weakest part of the article but fortunately this does not affect the main purpose : the development of an axiomatisation. I will start with a discussion of the shortcomings of paragraph 2 of the article. Quotations will be placed between asterisks.

2. Critique of Classical Theories;

* L.1 Clausius' Law ....

L.2 Kelvin's Law ....

L.3 Caratheodory's Law : In any neighbourhood of any state s of an isolated thermodynamic system there exist states which cannot be reached from s by any possible processes. *

Cooper's formulation of Caratheodory's Law is identical with the original formulation if 'possible processes of an isolated system' may be identified with 'adiabatically possible processes.'

(a) There exists a real valued empirical temperature function (s), which is an equivalence relation.

*(b) The work done in a small quasistatic adiabatic change of a thermodynamic system is given by a differential form d Q. *

The specification 'quasistatic' in (b), which is also present in Caratheodory's article, seems unnecessary, and is presumably a mistake in the latter article. The notation d Q for 'work done' and d W for 'heat transferred' is disturbingly unfortunate. The concept of 'heat' is not defined in Cooper's account. The Caratheodory approach attempts a mechanical definition of heat on the basis of the first law. A form of the first law deserves to be mentioned under the important assumptions, (a), (b), leading to the entropy law:

*L.4 The Entropy Law. There is a universal function of empirical temperature $T(\tau)$, which is itself an empirical temperature and a function of state φ such that $d\,W = T\,d\varphi$. *

*A. Either L.1 or L.2 implies L.3. *

The proof that both the Clausius Law and the Kelvin Law imply Caratheodory's Law leads to the conclusion that given any state, of a compound system, there exist states arbitrarily near it which are inaccessible, under complete isolation, because of Clausius' Law or Kelvin's Law. This does not imply Caratheodory's Law, which considers adiabatically possible processes of a bigger class of systems which contains in particular also simple systems.

*B. L.3 implies that there is a universal function $g(\tau)$ of empirical temperature such that $d\,W = g(\tau)\,d\varphi$ for some function φ. *

* L.1 implies L.4
L.2 does not imply L.4 : L.2 is equivalent to L.3 with the addition of the assumption that the work done in any isothermal change is non zero. *

It seems that in the proof of these theorems the concepts of 'heat' and 'work' are mixed up. For a Carnot cycle between temperatures τ_1 and τ_2, $\tau_2 < \tau_1$ is stated: * It follows from B that if the change of entropy at τ_1 is $\Delta\varphi$ and at τ_2 is $- \Delta\varphi$ then the quantities of work (sic!) done at these temperatures are $g(\tau_1)\ \Delta\varphi$ and $- g(\tau_2)\ \Delta\varphi$, respectively, * If the use of the term 'work' in this context is not a mistake, then the 'cycling' system must be an ideal gas (along an isotherm: $d\ Q = -d\ W$) and thus the existence of ideal gases are presupposed, which is denied in the sequel of the argument. The proof continues: *Clausius' Law asserts that these cannot be equal: otherwise the cycle would result only in a transfer of heat; * This assumes that the work done along the two adiabatic curve pieces of the cycle cancel each other, which is, in general, not true. The traditional argument would be: if the <u>heat</u> taken up at temperature τ_1 is equal to the <u>heat</u> given off at τ_2, then $\oint(d\ Q + d\ W) \neq o$ along the cycle, which contradicts the first law: thus $g(\tau_1) \neq g(\tau_2)$.

The proofs of this paragraph are not acceptable in the given form and the question of the relationships between the different forms of the second law needs a further examination. This is, however, outside the scope of this chapter.

3. <u>Accessibility conditions and Entropy functions</u>

* S.1 The state space S of a thermodynamic system $\mathcal{G}$ is a separable topological space. *

The hypothesis of separability needs additional specification to be physically fruitful. This specification is given in the definition of a simple system (p.187). In C P T the axiom is an implication of the supposed 'measurability' of the systems, with 'meters' for which real valued internal energy and deformation coordinate functions could be defined on the basis of topological properties expressed in $(\mathcal{S}, \mathcal{T}_{\underline{\to}})$. (see C P T 3.16, 3.20, 4.3).

* Acc 1. $\underline{\to}$ is a linear preorder in S *

* L.3

If the interval topology $(S, \mathcal{T}_<)$ is chosen, such that $s_1 < s_2$ iff $s_1 \rightarrow s_2$, then L.3´ is trivial, except the implication that there do not exist 'first states'. The author has presumably in mind that the state space is determined through 'mechanical coordinates' e.g. the internal energy U and deformation coordinates X_k in case of 'simple systems'. I agree that for the restricted purposes of this paragraph only the separability of the phase space is needed.

* Acc. 2 If $s_1 \rightarrow s_2$ then there are neighbourhoods $N(s_1)$ and $N(s_2)$ of s_1, s_2, respectively, such that if $s \in N(s_1)$ then $s_1 \rightarrow s_2$ and if $s \in N(s_2)$ then $s_1 \rightarrow s$ *

Again, some specification of the topology of the state space is desirable to give the axiom physical significance and to exclude a choice $(S, \mathcal{T}_<)$ which makes the axiom trivial. The counterpart of the axiom in C P T is the assumption that $S(U, \ldots X_k, \ldots)$ is continuous at all occupied points of phase space (C P T 4.26).

4. Composition of systems : Additivity of Entropy

* Systems $\mathcal{S}^1$ and $\mathcal{S}^2$ will be said to be isomorphic if there is a one-one map of the state space of $\mathcal{S}^1$ onto that of $\mathcal{S}^2$ which is a homeomorphism and preserves all thermodynamic relations. *

The corresponding term in C P T is 'identical'. Whether systems; which differ only in extent are excluded from being isomorphic is not certain: in this case there exists a one-one map which associates 'similar' states (see C P T 0 (2)).

* A system $\mathcal{S}$ is called the composition of systems $\mathcal{S}^1, \mathcal{S}^2, \ldots, \mathcal{S}^n$ and is written $\mathcal{S} = \{\mathcal{S}^1, \mathcal{S}^2, \ldots \mathcal{S}^n\}$ if there is a homeomorphism of the product space $S^1 \times S^2 \ldots . \times S^n$ onto the state phase of $\mathcal{S}$ such that if $\{s^1, s^2, \ldots, s^n\}$ is the state corresponding to $(s^1, s^2, \ldots, s^n)$ then:

Int (a) $\{s^1, s^2, \ldots s^{r-1}, s_1{}^r, s^{r+1}, \ldots, s^n\} \rightarrow \{s^1, s^2, \ldots s^{r-1}, s_2{}^r, s^{r+1}, \ldots, s^n\}$ if and only if $s_1{}^r \rightarrow s_2{}^r$.

Int (b) If $\mathfrak{S}^p$ is isomorphic with $\mathfrak{S}^q$ the state derived from $(s^1, \ldots, s^p, \ldots, s^q, \ldots)$ by permutation of s^p and s^q is reversibly accessible from it.

Int (c) If A, B are complementary subsets of {1, 2, ...n} and $\mathfrak{S}^A$ is the composition $\{\mathfrak{S}^{a_1}, \mathfrak{S}^{a_2}, \ldots \mathfrak{S}^{a_p}\}$ and $\mathfrak{S}^B$ is $\{\mathfrak{S}^{b_1}, \ldots \mathfrak{S}^{b_{n-p}}\}$ then $\mathfrak{S} = \{\mathfrak{S}^A, \mathfrak{S}^B\}$ *

The relationships between this definition and the 'metrization axiom' in C P T (2.9) are straightforward. C P T 2.9 iii: if $z' \ldots z' \Rightarrow z'' \ldots z''$, then $z' \Rightarrow z''$ has no counterpart in the above definition, but this property is used in Cooper's proofs as well. It might be necessary to consider iii also as a defining property of compositions in Cooper's sense.

* Int 1. There is a class of elementary systems, which have simply connected state spaces. This class contains at least four systems isomorphic to a system $\mathfrak{S}$. The composition of any four elementary systems exists. *

The elementary systems, with connected state spaces correspond in C P T with an entropymeter, which can be duplicated. The meaning of the term 'connected' in this axiom becomes clear through its consequence: the existence of (empirical) entropy functions whose values form an interval of the reals. Connectedness is thus closely related to the first property of the entropymeter in C P T: $(\mathcal{S}_\sigma, \mathcal{G}) \simeq (R', <)$ (C P T 2.17).

I do not feel the need to avoid the assumption that the composition of any two thermodynamic systems exists. In a discussion of the relations between mathematical foundations and physical content of a theory a distinction between mathematical existence and physical existence might be helpful. It is also not clear to me in what sense the assumption leads to an infinite regress.

* **Theorem 2.** Under the assumptions of Acc 1-3 and Int 1, there is a function φ which is defined on the state spaces of all elementary systems, which is entropy function for any one system, and which is such that the function $\varphi(s^1) + \varphi(s^2)$ of $\{s^1, s^2\}$ is an entropy function for $\{\mathfrak{S}^1, \mathfrak{S}^2\}$ *

This theorem is identical with C P T 2.19 and 2.16. A part of the proof in C P T is borrowed from Cooper's proof: Lemma 2 and its proof is adopted without alterations, the extension from states with dyadic entropy values to all states is carried out more explicitly. The definition of an entropy function for arbitrary elementary systems and the proof of the extensivity of this function is closely followed.

The essential difference between the proof of theorem 2 and the approach of C P T is caused by the impossibility of using Caratheodory's principle (L.3′) and Buchdahl and Greve's continuity assumption (Acc 2) before a phase space, preferably $(U, \ldots X_k \ldots)$ is specified. In C P T preference is given to the definition of extensive internal energy and deformation coordinate functions and extensive entropy functions, on the basis of ad hoc suppositions with respect to the properties of the meters; next the maximum entropy principle is introduced together with continuity assumptions which imply Caratheodory's principle and the associated continuity assumption. This choice has consequences for the proof of Theorem 2 in C P T: the first part of lemma 1 cannot be proved and is introduced as a defining property of the entropymeter. Lemma 3 is replaced through the 'calibration property' of the entropymeter. In the further development of C P T the properties of the entropymeter may be considered as a consequence of connectedness (global) of the phase space of the entropymeter.

5. Internal Energy and Temperature

The content of this paragraph can be characterised as an attempt to define the concepts 'adiabatic interaction', 'adiabatic isolation,' 'internal energy,' 'diathermal interaction' and 'thermal equilibrium' in terms of the 'ground theories.' This conforms to the traditional approach which tries to define thermodynamical concepts on the basis of mechanics, e.g. the 'mechanical definition of heat?

In C P T a quite different attitude is explained: adiabatic isolation and diathermal connection are primitive terms of the theory. It is a certain preorder relation and a certain equivalence relation on the set of all states. A definition may be attempted on a level different from the formal theory, but a satisfactory reduction to mechanics (or more general: 'ground theories') does not exist in my opinion. Similar remarks apply to 'internal energy.' It is in my opinion an irreducible thermodynamic concept. A definition of internal energy on the basis of the first law, as in E 1, introduces the concept of an adiabatically isolated system, and the latter concept cannot be explained in purely mechanical terms. In C P T an internal energy function is defined on the basis of a primitive 'energetic isolation.' The extensivity (additivity) of a certain constructed energy function can be proved, and this procedure which is closely related to traditional calorimetry, seems to me preferable to the axiomatic assertion of the additivity of the internal energy in E 2.

* Temp 1. There exists a real valued function $\theta(s)$ defined for all states of all simple thermodynamic systems, which is such that two states s^1 and s^2 of any two systems are in equilibrium if and only if $\theta(s^1) = \theta(s^2)$. For a fixed configuration x, $\theta(x, E)$ is a strictly increasing function of E. *

This axiom has three parts: thermal equilibrium is an equivalence relation on the set of all pairs of states of simple thermodynamic systems; there exists a map from the set of equivalence classes of equal temperature onto a subset of the reals; $\theta(x, E)$ is a strictly increasing function of E at constant x. These three statements are part of traditional thermodynamics in a Kelvin-Clausius or Caratheodory approach. I agree with Cooper that the proof of the existence of a real valued empirical temperature function from a Zeroth law cannot be carried out without additional suppositions which make the profitability of the proof questionable.

In C P T the first part of Temp 1 is present as an axiom (C P T 1.2). The introduction of an empirical temperature function appears not to be necessary; the maximum entropy principle and associated continuity assumptions imply the existence of an absolute temperature function $1/T \equiv (\partial S/\partial E)_x$ and the convexity upwards of $S(E)_x$, from which follows that $T(x, E)$ is a strictly increasing function of E at constant x.

6. Absolute Temperature

* Acc 4. If $\mathcal{S}^1$ and $\mathcal{S}^2$ are any two systems which are in thermal equilibrium in s_o^1, s_o^2, then no other state $\{s^1, s^2\}$ of the composed system $\{\mathcal{S}^1, \mathcal{S}^2\}$ in which each system is in the same configuration is accessible from the state $\{s_o^1, s_o^2\}$. *

This axiom can be regarded as the maximum entropy principle in a global form, restricted to a subspace of constant 'configuration' (only exchange of the energy between the part systems is permitted).

The proof that $S(E)_x$ is convex upwards and that $(\partial S^1/\partial E^1)_{x^1} = (\partial S^2/\partial E^2)_{x^2}$ if $\theta(s^1) = \theta(s^2)$ is standard. That the left hand and right hand derivatives are equal is not proved convincingly: perhaps continuity of $\theta(E, x)$ is supposed.

Similar procedures are followed in C P T. The maximum entropy principle in C P T is however more general than Acc 4, and consequently more general conclusions can be drawn. The equality of the left and right-hand derivatives is asserted axiomatically.

7. Absolute Temperature as an Integrating Factor

This last paragraph of Cooper's article may be considered as an analysis of the topological presuppositions of the Caratheodory approach. The definition of a 'smooth' system explains in precise terms the existence of isotherms (b) and adiabatic curves or hypersurfaces (d) and their intersection everywhere (e) and the differentiability of the generalised forces with respect to the internal energy E (c). The property (a) can be shortened to "G is simple", after the definition of 'simple' in paragraph 5. One step of the argument of this paragraph requires a further analysis. It seems to be tacitly assumed that there exists an integrating factor for $d E + p \, dt$ which is a function of θ only. A derivation of this statement is an important part of the Caratheodory approach. A discussion of this derivation and a reconstruction of the argument within the framework of Cooper's axiomatisation would be desirable.