

## *Introductory note to 1906*

Jos Uffink

Josiah Willard Gibbs's book *1902* provided an approach to the foundations of statistical physics that differs in many respects from Boltzmann's. Perhaps the foremost difference, at least for present concerns, is that Gibbs aimed to provide what he called a "rational" foundation to the statistical approach, while refraining from "solving the mysteries of nature". Indeed, Gibbs's book reads like an exposition of the rigorous theorems of a theory that he baptized "statistical mechanics", assuming only the most general assumptions needed for such an account.<sup>1</sup> For example, the assumption that the systems under consideration are actually composed of a finite number of molecules is only introduced in the final chapter. And whenever Gibbs discusses the consequences that this theory would have for the reduction of thermodynamics, these discussions are embedded in thoughtful reservations, pointing out that they merely suggest certain analogies between the two theories—never a bold statement that he had shown that a law of thermodynamics could rigorously be reduced to a theorem in another theory, as Boltzmann did on various occasions. For example, Gibbs emphasized that the analogies in question are far from unique. His argument was rather that the existence of such analogies could make it seem plausible that thermodynamical relations are not incompatible with the statistical-mechanical approach. Roughly speaking, Gibbs was a cautious theoretical physicist who kept a keen eye on mathematical rigor, whereas Boltzmann, at times, boldly announced claims that overstated what he had actually proved. It seems to me that Zermelo must have been much more sympathetic to Gibbs's approach than to Boltzmann's, since Gibbs's goal seems to coincide with the goal stated in Zermelo's *1896a*, i.e. to state clearly what can and cannot be proved. More extensive discus-

---

<sup>1</sup> The distinction between "statistical mechanics" proper and the more general concept of "statistical physics" is somewhat subtle. It is typical for statistical mechanics that one assumes a mechanical phase space for the entire system considered (whether this be a gas, liquid or another macroscopic body), together with the Hamiltonian equations of motion, and a probability measure defined on this phase space. This stands in contrast to other approaches within statistical physics, for example the kinetic theory of gases, where probability considerations are applied on the level of a single molecule. Although Gibbs invented the term "statistical mechanics" and developed this theory systematically, this is not to say that he originated the approach. Indeed, several of Boltzmann's papers took the same approach, and so did Zermelo. However, Boltzmann never clearly distinguished the cases where he was pursuing statistical mechanics, as we now use this term, from cases where he was pursuing kinetic theory or some other set of assumptions within statistical physics.

sions of the difference between the Boltzmannian and Gibbsian approaches are offered by *Frigg 2008*.

Zermelo produced the first German translation of Gibbs's book. I do not know whether this translation was commissioned by the publisher or whether it was undertaken on Zermelo's own initiative. In the latter case, one would suppose that the effort expended would indicate a significant affinity that Zermelo must have felt towards Gibbs's work. But in any case, Zermelo's appreciation for Gibbs is evident from the words he wrote in the preface to the translation (Zermelo in *Gibbs 1905*, iii–iv):

This is the first effort to develop rigorously and on a secure mathematical basis the statistical and probabilistic considerations in mechanics which are indispensable in various areas of physics, especially in the kinetic theory of gases. As such, this work is of paramount interest and of permanent value, even if not all the derivations by the author should turn out to be tenable. Which objections I have to make, from my side, in particular against the considerations in Chapter 12, in as far as these attribute to mechanical systems a tendency towards a state of statistical equilibrium, and on this basis, a complete analogy with thermodynamical systems, I intend to explicate on another occasion. [...]

May this work in its new form contribute to raise the interest for and understanding of statistical mechanical problems amongst German physicists and mathematicians.<sup>2</sup>

The “another occasion” to which this quote refers is Zermelo's own review *1906* of Gibbs's book. After a clear statement of the problems and purpose of statistical mechanics as set out by Gibbs, Zermelo argues that similar arguments had been used earlier in the theory of gases, but that their application “was not always unobjectionable” (p. 234). This, of course, is a veiled reference to Boltzmann. By contrast, Gibbs is praised for developing

---

<sup>2</sup> “Es ist der erste Versuch, die statistischen und die Wahrscheinlichkeits-Betrachtungen in der Mechanik, wie sie auf verschiedenen Gebieten der Physik, namentlich aber in der kinetischen Gastheorie unentbehrlich sind, unabhängig von ihrem Anwendungsgebiet auf sicherer Grundlage mathematisch streng zu entwickeln, und ist als solcher von hervorragendem Interesse und bleibender Bedeutung, auch wenn sich nicht alle Deduktionen des Verfassers als haltbar erweisen sollten. Welche Einwendungen ich meinerseits namentlich gegen die Betrachtungen des zwölften Kapitels zu machen habe, soweit sie den mechanischen Systemen eine Tendenz nach einem Zustande statistischen Gleichgewichts und auf Grund dessen eine vollständige Analogie mit den thermodynamischen Systemen im Sinne des zweiten Wärmesatzes zuschreiben, gedenke ich demnächst an anderer Stelle auszuführen. [...]

Möge das Werk in seiner neuen Form dazu beitragen, das Interesse und Verständnis für statistisch-mechanische Probleme bei deutschen Physikern und Mathematikern zu fördern.”

the foundations of the subject, independently of the intended domain of application, on a rigorous mathematical basis. Nevertheless, Zermelo has two main objections against Gibbs's work. The first has to do with the analogies that Gibbs aims to point out between relations holding in statistical mechanics and those in thermodynamics. The second deals with his treatment of irreversible processes.

In Gibbs's approach, as is well known, one considers a probability measure, say  $\rho$ , which may be represented by a density function  $P$  on the phase space  $\Gamma$ , such that

$$\rho(A) = \int_A P(x) d\mu(x)$$

represents the probability of finding the state of the system at some given initial time in a measurable set  $A \subseteq \Gamma$ . (Zermelo writes  $dp_1 \cdots dq_n$  for  $d\mu(x)$ , the Lebesgue measure on  $\Gamma$  in the canonical coordinates.)

This probability is physically interpreted by reference to an *ensemble*, i.e. one imagines a great number  $N$  of replicas of the system under consideration with arbitrary initial states and supposes that the relative number of systems that have their initial state in some measurable region  $A$  converges, as  $N \rightarrow \infty$ , to  $\int_A P(x) d\mu(x)$ .

The average or expected value of any phase space function  $f(x)$  over this ensemble is then given by

$$\bar{f} = \int_{\Gamma} f(x) P(x) d\mu(x). \quad (1)$$

In general, the probability density function  $P$  will change with time, in accordance with the flow generated by the Hamiltonian equations of motion. But in certain cases, it can be invariant under time evolution. These cases Gibbs called *statistical equilibrium*. In general, this will happen if  $P$  depends only on phase space functions that are conserved by the Hamiltonian flow. And since the Hamiltonian  $H$  itself is always conserved under the Hamiltonian flow, any probability density that depends only on  $H$  will be in statistical equilibrium.

Gibbs devoted special attention to a particular family of probability density functions  $P$  in statistical equilibrium which he called *canonical*, which in modern notation would be written as:

$$P(x) = \frac{1}{Z(\beta)} e^{-\beta H(x)} \quad (2)$$

where  $H$  is the Hamiltonian,  $\beta$  an arbitrary parameter related to the temperature, and  $Z(\beta)$  a normalization factor ensuring that  $P$  is indeed a probability density, i.e.  $Z(\beta) := \int_{\Gamma} e^{-\beta H(x)} d\mu(x)$ . In Gibbs's own notation, also used by Zermelo, this corresponds to

$$P(x) = e^{\frac{\psi - \epsilon}{\theta}}, \quad (3)$$

where  $\epsilon$  denotes the Hamiltonian  $H$ ,  $\theta = \frac{1}{\beta}$ , and  $\psi = \ln Z(\beta)$ .

For these canonical distributions (or the ensembles they represent), Gibbs argued that one could obtain relations which were analogous to results in equilibrium thermodynamics if we consider  $\theta$  analogous to the system's absolute temperature, and

$$\bar{\eta} = \int P(x) \ln P(x) d\mu(x) \quad (4)$$

as analogous to (minus) its entropy.

In particular, assuming that the Hamiltonian depends smoothly on some parameters  $a_i$  ( $i = 1, \dots, n$ ), and if  $A_i := \frac{\partial A}{\partial a_i}$ , Gibbs shows that

$$d\bar{H} = -\theta d\bar{\eta} - \sum_i \bar{A}_i da_i \quad (5)$$

which, at least typographically, is similar to Gibbs's own (1875) formulation of the fundamental equation of thermodynamics, which (in the special case of fluids) reads

$$dE = TdS - pdV. \quad (6)$$

So, if we consider  $\theta$  to be analogous to temperature,  $-\bar{\eta}$  to the entropy  $S$ , take the Hamiltonian to depend only on a single parameter, i.e. the volume  $V$ , and the pressure  $p$  to be analogous to  $\bar{\partial H} / \bar{\partial V}$ , we recover this fundamental equation of thermodynamics in statistical mechanics, at least for this special case of a fluid described by a canonical ensemble.

However, Zermelo noticed serious problems with this proposed analogy (p. 237):

But this analogy remains entirely superficial and cannot clarify the mechanical character of the thermodynamical equation. For the magnitude  $\bar{\eta}$  [...] is not at all a property of the individual mechanical systems.

Indeed, while the thermodynamical relation (6) is supposed to relate the differential change of energy  $dE$  of an individual system to the changes in its entropy  $S$  and volume  $V$ , some quantities appearing in the suggested analogy (5), in particular  $\theta$  and  $\bar{\eta}$ , are only defined by recourse to the envisaged ensemble. But any individual system can be regarded as being part of any ensemble whatsoever, and properties of the ensemble need not reflect properties of an individual system.

This remark by Zermelo is a penetrating one, and points out a deep problem in Gibbs's approach that still remains unsolved today. The point here is not just that there is a distinction between an individual system and an ensemble. For example, the energy of an individual system, as characterized by statistical mechanics, would be  $H(x)$ , where  $x \in \Gamma$  is the state of the individual system. The putative analogy (5), instead, involves the ensemble average  $\bar{H}$ , which might very well be very different from the energy value  $H(x)$

of the individual system. In general, an individual system need not resemble an average system. Similar remarks could be made about the pressure of the individual system  $\partial H(x)/\partial V$  and the ensemble average  $\overline{\partial H/\partial V}$ .

But *this* point was considered by Gibbs: He showed that for systems with many constituents, the canonical probability density, considered as a function of energy, becomes increasingly peaked, and that the standard deviation in energy will approach zero if the number of constituents goes to infinity. So, if the system is large enough, the probability that (say)  $H(x)$  deviates more than a tiny amount from  $\overline{H}$  is negligible.

However, the problem that Zermelo pointed to is much deeper: His point, as I read it, is that  $\theta$ , the putative candidate for temperature, is not even a function on phase space.  $\theta$  is just a parameter in the canonical probability density (3) on  $\Gamma$ , and so for an individual system in a state  $x \in \Gamma$ , there is no way of determining what its value is. The quantity  $\theta$  does not even refer to a property that the individual system might have. Something similar holds for  $\eta(x)$  and  $\overline{\eta}$ . Here, perhaps, one might say that  $\eta(x) = \ln P(x)$  actually does define a function on  $\Gamma$  and that each individual system in state  $x$  will have the corresponding value  $\eta(x)$ . But the phase space function  $\eta$  is not fixed by the mechanical constitution of the system, like the Hamiltonian  $H$  or the pressure  $\partial H/\partial V$ . It is only defined by recourse to the ensemble from which the individual system is taken to be a member. But any individual system can be considered as a member of many different ensembles, even if we just focus on canonical ensembles. Consequently, we cannot determine the value of  $\eta(x)$  for an individual system in state  $x$  unless we are told which ensemble it is supposed to belong to, so that we can specify the form of  $P$  in  $\eta(x) := \ln P(x)$ . The value  $\eta(x)$  is just not a mechanical property of an individual system in state  $x$ , i.e. a property that can be ascertained independently of the choice of ensemble. And, of course, a demonstration that, for canonical ensembles of systems whose number of constituents increases to infinity, the standard deviation in  $\eta$  goes to zero, as Gibbs provided, does nothing to alleviate this point.

Of course, I admit that it is not completely clear that Zermelo's briefly stated objection actually meant to point out this particular problem. However, if the above reading of what Zermelo was pointing out is even roughly correct, it shows a remarkably deep insight by Zermelo into the foundational problems of Gibbs's approach. I would not know of any other author in the early 20th century that expressed awareness of this problem. Whether Gibbs himself was aware of this problem is quite unclear. He expressed, as far as I can tell, no concerns about this issue, but one could imagine it might be one of the reasons why he was so cautious in calling the relationship between (6) and (5) a mere "analogy".

Gibbs's discussion of this topic might well be compared to Boltzmann's discussion in *1871b* of a similar issue, to which Gibbs refers several times in his book. Boltzmann, in effect, had already introduced in this paper what we

today (following Gibbs) call the canonical ensemble and applied it to study infinitesimal changes in the expected value of energy, to obtain a relation that was typographically identical to a relation in thermodynamics, similar to (6).<sup>3</sup> However, as mentioned before, the striking difference between Gibbs's discussion of this issue and *Boltzmann 1871b* is that Boltzmann presented his result boldly as an "analytical proof of the second law of thermodynamics", whereas Gibbs presented his result as merely establishing an "analogy".

The second objection that Zermelo raised in his *1906* is a return to the topic of his 1896–7 dispute with Boltzmann. Can statistical mechanics, rigorously developed, ever entail a description of irreversible processes? Gibbs discusses this point in his Chapter 12.

In this chapter, Gibbs provides a discussion of Poincaré's recurrence theorem and the objections it entails to the monotonic increase of entropy. Whether he learned about this theorem from studying Poincaré, or by reading the debate between Zermelo and Boltzmann, is unclear, since he gives no references. Gibbs reaches from this discussion more or less the same conclusions as Zermelo did about the feasibility of incorporating an irreversible approach to equilibrium in statistical mechanics, i.e., "we find by this method no approach to statistical equilibrium in the course of time (*1902*, 144)."

However, Gibbs did not draw the same devastating conclusions that Zermelo (*1896a*) did about the prospects of statistical mechanics to deal with irreversibility. Instead, he warned the reader to withhold judgment: "Yet we must here exercise extreme caution (*ibid.*)" Gibbs's discussion continues by pointing to an analogy between the mechanical motion in phase space and the stirring of two differently colored dyes, conceived of as incompressible fluids. When two such dyes, say black and white, are mixed by an external stirring device, the total volumes occupied by the black and white dyes remain invariant. Even so, these volumes can become so deformed and intermingled by the mixing operation that, if one does not look too closely, one may regard the resulting mixture as being uniformly gray.

Gibbs suggests that something similar might apply to mechanical motion in phase space (in this case, however, without external intervention). Thus, even if we start out with a probability density function on the phase space  $\Gamma$  which does not represent statistical equilibrium (say it is zero outside some region  $A \subseteq \Gamma$  and a constant inside that region, and the constant and zero are taken as analogous to black and white), the autonomous evolution of that region under Hamiltonian mechanics might be such that this region develops into ever more branching and fibrillating tentacles that will in some sense cover the entire phase space uniformly, at least if we do not look too closely.

Gibbs showed that the mechanical evolution will necessarily conserve the value of  $\bar{\eta}$  for any non-equilibrium probability density. Thus, at first sight,

---

<sup>3</sup> Actually, in *1971b* Boltzmann does not propose a concrete expression analogous to entropy, as Gibbs does in *1902*.

the expression he intended to serve as an analogy of thermodynamical entropy,  $-\bar{\eta}$ , could not increase in the course of time, even for a probability density that did not correspond to statistical equilibrium. But he did not conclude that his theory implies that entropy must be conserved under any Hamiltonian mechanical evolution. Instead, Gibbs argued that if the above picture applies, the probability density function that results after a long time in this process of fibrillation might be well approximated by a uniform, gray, distribution, and for this approximating uniform probability density, the expression  $\bar{\eta}$  would actually be less than that of the probability density we started out with, corresponding to an increase in thermodynamical entropy. Accordingly, Gibbs diagnosed the problem of representing irreversible processes in statistical mechanics as a problem of the interchange of the order of two different limits in the representation of physical systems: one limit of taking a finer and finer coarse-graining, and the other limit of going to infinite time in the evolution of the density function.

Gibbs's suggestions here have been taken up by later authors in an attempt to fill in the details. The first ingredient is to distinguish between two kinds of entropy: "coarse-grained" versus "fine-grained", a distinction first proposed by Poincaré (1906) (as *entropie grossiere* and *entropie fine*). If we take the fine-grained Gibbs entropy as generally defined by (4),

$$-\bar{\eta}(P) = - \int P(x) \ln P(x) d\mu(x),$$

the coarse-graining is introduced as an averaging of  $P$  over some partition of phase space:  $\Gamma = \omega_1 \cup \dots \cup \omega_m$  (with  $\omega_i \cap \omega_j = \emptyset$  if  $i \neq j$ ). The coarse-grained version of  $P$  is given by

$$\text{CGP}(x) = \sum_{i=1}^m \frac{1_{\omega_i}(x)}{\mu(\omega_i)} \int_{\omega_i} P(x) d\mu(x)$$

where  $1_{\omega}$  is the characteristic function of the set  $\omega$ ; and the coarse-grained entropy of  $P$  is then defined as the Gibbs entropy of the coarse-grained version of  $P$ :

$$\Sigma[P] = -\bar{\eta}(\text{CGP}).$$

One can show that, while the fine-grained entropy  $-\bar{\eta}(P_t)$  is stationary under time evolution, the coarse-grained entropy can change its value under this evolution.

The other ingredient, stressed in particular in the 1940s by Nikolai Sergeevich Krylov (1979), is to introduce as an additional assumption that dynamical systems have the property known as "mixing". In the context of the theory of dynamical systems, the mixing property is the requirement that, for any two measurable subsets  $A, B \subseteq \Gamma$ :

$$\lim_{t \rightarrow \infty} \mu(T_t A \cap B) = \mu(A)\mu(B). \quad (7)$$

This seems a plausible formalization of the dynamical property that Gibbs had in mind when he compared the evolution to the stirring of dyes.

However, when it comes to answering the question whether these ingredients would suffice to provide a statistical-mechanical analogue of the thermodynamical approach to equilibrium, the situation is somewhat diffuse. Thus, although it can be shown that if the mixing assumption holds, and for an appropriate choice of the partition, the coarse-grained entropy  $\Sigma(P_t)$  of any continuous non-equilibrium probability density  $P_t$  will approach, for  $t \rightarrow \infty$ , the value of the coarse-grained entropy of statistical equilibrium, this approach need not be monotonic. Moreover, the same is true for the limit  $t \rightarrow -\infty$ .

To return from these later developments to Zermelo's discussion, it seems that Zermelo dismisses Gibbs's considerations almost immediately as being ad hoc. He argues that there is no reason to suppose that a system should have the mixing property. In a sense, Zermelo is right here. Yet to me his dismissal also seems a bit too easy. Of course, it is true if there are many dynamical systems, even Hamiltonian ones, that do not have the mixing property. So, if one takes the purpose of statistical mechanics to be the systematic and general exposition of rigorous results pertaining to all Hamiltonian systems without any regard to possible applications in the kinetic theory of gases, as Zermelo seems to understand Gibbs's aims to be, then the appeal to this mixing property might seem rather ad hoc.

However, there are also many examples of dynamical systems that do have the mixing property. In other words, there is nothing in the general formalism of statistical mechanics that makes this assumption incompatible. And so, it seems a perfectly viable strategy to argue that when it comes to explaining the observed tendency towards equilibrium—an issue that only arises in special applications like in the kinetic theory of gases—to appeal to the conjecture that those systems that display this tendency may happen to possess an additional property, in order to explain this particular type of behavior.<sup>4</sup>

It is another matter whether the strategy is also successful. Here, another objection raised by Zermelo is more to the point: even if the mixing property were to hold, it would not break the time-reversal invariance of the theory, and imply not only a tendency towards statistical equilibrium towards the future (i.e. the limit  $t \rightarrow \infty$ ) but also towards the past,  $t \rightarrow -\infty$ . Hence, as Zermelo puts it, the property gives no basis to conclude that if we take two arbitrary times  $t'' > t'$  that the Gibbs entropy at one instant should be less than at the other.

---

<sup>4</sup> The question whether the mixing property holds for the kind of systems of interest to the intended application, e.g. the hard spheres model, turns out to be extremely tedious (*Sinai and Chernov 1987, Szász 1996, Simányi and Szász 1999*).



Zermelo's papers on the Gibbs approach to the foundations of statistical mechanics are not well-known, but I hope to have made clear that he made a major contribution in making Gibbs's work accessible to the German-speaking community and, in his critical review of 1906, made a couple of

## Besprechung von *Gibbs 1902* und *Gibbs 1905*

1906

233 Seit den frühesten Anfängen der Naturwissenschaft ist es ein Lieblingsgedanke der Forscher gewesen, die Mannigfaltigkeit der Naturvorgänge auf eine gemeinsame Grundform, auf die Bewegung ihrer kleinsten Teilchen, zurückzuführen. Namentlich für eine exakte quantitative Naturerkenntnis schien eine solche Zurückführung aussichtsvoll, nachdem durch *Galilei* und *Newton* die Grundlage für eine mathematische Theorie der Bewegungserscheinungen geschaffen war. Es ist bekannt, daß *Newton* selbst die Optik in Gestalt der Emissionstheorie als einen Zweig der Mechanik behandelt hat, und wenn auch später diese Theorie zugunsten der Undulationstheorie aufgegeben werden mußte, so verstand man doch auch die letztere zunächst rein mechanisch, indem man sich den Äther wie eine elastische Flüssigkeit schwingend vorstellte. Eine Neubelebung erfuhr die mechanistische Vorstellungsweise, als mit der Entdeckung des Energiegesetzes die erkannte Äquivalenz von Wärme und Arbeit die Vermutung nahe legte, daß die Wärme eine reine Bewegungserscheinung und die Temperatur eines Körpers durch die lebendige Kraft der unregelmäßig bewegten Moleküle zu messen sei. Aber nur in der kinetischen Gastheorie, deren Hypothesen so besonders einfach sind, ist eine solche Zurückführung der thermischen Vorgänge auf mechanische bis zu einem gewissen Grade gelungen, bei den festen Körpern und den Flüssigkeiten ist sie bisher noch in ziemlich unvollkommenen Anfängen stecken geblieben.

Die Schwierigkeiten, die sich der exakten Durchführung der Molekularhypothese entgegenstellen, sind von zweierlei Art. Einmal entsteht die physikalische Frage: wie sollen wir uns die Beschaffenheit der kleinsten materiellen Teile vorstellen und welche Wirkungsgesetze den zwischen ihnen wirkenden Kräften zuschreiben, um einen bestimmten Ansatz für die Rechnung zu gewinnen? Es ist klar, daß wir uns hier mit ziemlich rohen Annäherungen, mit unvollkommenen Bildern begnügen können, sofern nur zu erwarten steht, daß die hierbei begangenen Fehler ohne wesentlichen Einfluß auf die Gesamterscheinung sein werden. So begnügt man sich in der Gastheorie meist mit der Vorstellung elastischer Stöße zwischen einfach geformten festen Körpern in der Annahme, daß das wahre Wirkungsgesetz der zwischen den Molekülen wirkenden Kräfte bei der kurzen Dauer der Einwirkung nur unwesentlich in

penetrating remarks, long before anyone else, pointing out that Gibbs's approach, while keeping close to the standard of a rigorous mathematical framework, did not provide a fully satisfactory explanation of the thermodynamical analogies, or the increase of entropy in isolated systems.

---

## Review of *Gibbs 1902* and *Gibbs 1905*

1906

Since the earliest dawn of natural science researchers have been taken with the idea that the manifold of natural phenomena should be reduced to a common basic form, to the motion of their smallest particles. Such reduction seemed particularly promising for an exact quantitative knowledge of nature once *Galilei* and *Newton* had created the foundation for a mathematical theory of the phenomena of motion. *Newton* himself, as is well-known, had treated optics in the form of emission theory as a branch of mechanics. Even though it was later necessary to abandon this theory in favor of the undulation theory, understanding of the latter was purely mechanical at first as it was based on the idea that ether oscillates like an elastic liquid. The mechanistic conception was reanimated when the discovery of the law of energy led to the realization that heat and work are equivalent, which suggested that heat is but a phenomenon of motion and that the temperature of a body is to be measured as the living force of molecules in irregular motion. But only for the kinetic theory of gases, whose hypotheses are particularly simple, was it possible to reduce the thermal processes to mechanical ones to a certain extent, while similar attempts undertaken for solid bodies and for liquids have remained stalled in their early, inchoate beginnings.

The difficulties standing in the way of the rigorous implementation of the molecular hypothesis are of two different kinds. First, the physical question arises: How are we to conceive of the constitution of the smallest material parts, and which laws are we to attribute to the forces acting among them, in order to obtain a definitive starting point for calculation? It is evident that we can make do here with crude approximations and incomplete pictures, provided only that we can expect that the subsequent errors will have no substantive bearing on the phenomenon as a whole. Thus, in applying the theory of gases, we are mostly content to consider elastic collisions between simply formed solid bodies, assuming that the true laws of the forces acting between the molecules will not really be relevant considering the short

Betracht kommen werde. Aber jetzt beginnt eine weitere, die mathematische Aufgabe, auf Grund unserer physikalischen Hypothesen die zusammengesetzten Vorgänge zu berechnen, die in den gesamten Molekülsystemen, d. h. in den unserer Beobachtung zugänglichen Körpern unter den gemachten Annahmen stattfinden würden, damit wir in der Lage sind, die so berechneten Vorgänge mit der Erfahrung zu vergleichen und daraus Rückschlüsse auf die Berechtigung unserer Annahmen zu ziehen. Bedenkt man nun, welche außerordentlichen Schwierigkeiten schon der einfache Fall der drei nach dem *Newtonschen* Gesetze bewegten Himmelskörper den Astronomen bietet, so erscheint es fast vermessen, die Bewegungsvorgänge in einem aus vielen Billionen von Molekülen zusammengesetzten Systeme berechnen zu wollen. Doch auch hier bietet sich eine bedeutende Erleichterung. Es ist ja weder notwendig noch auch wünschenswert, den Weg jedes einzelnen Teilchens zu verfolgen, sondern es genügt, die in ihrer Gesamtheit geltenden Gesetzmäßigkeiten zu erforschen. Wir brauchen weder den genauen Anfangszustand zu kennen, der uns durch messende Beobachtung doch niemals gegeben werden kann, noch auch den genauen Endzustand zu berechnen. Vielmehr kommt es nur darauf an, aus einem nur in ganz rohen Umrissen bestimmten Anfangszustand auf das Eintreten eines ebenso roh bestimmten Endzustandes zu schließen und dabei von allen singulären und wegen ihrer relativen Kleinheit und Seltenheit in der großen Masse | verschwindenden Einzelvorgängen abzusehen. Man muß mit einem Worte die Mechanik der Moleküle nicht individuell sondern *statistisch* betreiben. Nicht die Bewegung eines einzelnen Teilchens ist zu bestimmen, sondern die Veränderung gewisser Funktionen, die von unzählig vielen solchen Teilchen abhängen, und auch hier nur die durchschnittliche oder wahrscheinliche Änderung für alle möglichen mit dem beobachtbaren physikalischen Anfangszustand verträglichen Positionen. Bei dieser statistischen Betrachtungsweise ist auch anzunehmen, daß unsere Hypothesen über Gestalt und Beschaffenheit der Moleküle und ihre Wirkungsgesetze nur teilweise zum Ausdruck kommen und daß speziellere Vorstellungen für die Beantwortung vieler Fragen sich als unnötig erweisen. Freilich müssen Grundlagen und Methoden dieser „statistischen Mechanik“ zuerst bekannt sein, es muß möglich sein, ohne willkürliche und zweifelhafte Hypothesen die Massenerscheinungen in mechanischen Systemen mit sehr viel Freiheitsgraden unter Beschränkung auf eine plausible Wahrscheinlichkeit ebenso sicher zu bestimmen wie vermöge der reinen Mechanik die Vorgänge in einem einzelnen System mit einer beschränkten Zahl von Freiheitsgraden.

Solcher Methoden hat man sich natürlich in der Gastheorie immer bedient. Aber die Art und Weise, wie man die in allen Richtungen durcheinander fliegenden Moleküle durch Gruppen von mittlerer Beschaffenheit ersetzte, wie man die Verteilung der Geschwindigkeiten nach dem *Maxwellschen* Gesetze zu begründen und die *Wahrscheinlichkeitsrechnung* auf die mittlere Anzahl und den Effekt der Zusammenstöße anzuwenden suchte, war nicht immer einwandfrei. Auch bezogen sich alle diese Methoden meist auf den speziellen nur in der Gastheorie vorliegenden Fall eines Systems von geradlinig beweg-

duration of the action. But now a further, mathematical task awaits us: to calculate on the basis of our physical hypotheses the composite processes that would arise in the whole systems of molecules, i. e., in the bodies accessible to our observation, given the assumptions made, in order to be able to compare the processes thus calculated with experience and to draw conclusions from this about the extent to which our assumptions are justified. Considering what extraordinary difficulties even the simple case of three celestial bodies moving in accordance with *Newton's* laws creates for astronomers, it appears almost presumptuous to try to calculate the processes of motion in a system composed of many trillions of molecules. But here, too, we are afforded significant relief. For it is neither necessary nor desirable to chart the path of each individual particle but, rather, sufficient to investigate the lawful regularities obtaining in their totality. We do not need to know the precise initial state, with which observation by measurement could never acquaint us anyway. Nor do we need to calculate the precise final state. Instead, all that matters is to deduce from a particular initial state given only in very rough outlines a final state determined in a similarly rough fashion while ignoring the singular individual processes that disappear in the great mass due to their relative small scale and their rarity. In a word, we must pursue the mechanics of the molecules not individually but *statistically*. Instead of determining the motion of individual particles, we must determine the changes of certain functions that depend on innumerable many such particles, and more specifically only the average or probable change for all possible positions compatible with the observable physical initial states. It is to be assumed that, in this statistical approach, our hypotheses about the form and constitution of the molecules and the laws governing them find only partial expression and that many questions need no more specific conceptions to receive an answer. Of course, we must first know what the basic ideas and methods of this "statistical mechanics" are. Restricting ourselves to a plausible probability, we must be able to determine, without relying on arbitrary and dubious hypotheses, the mass phenomena in mechanical systems with a great number of degrees of freedom with as great a certainty as we determine the processes in an individual system with a restricted number of degrees of freedom by means of pure mechanics.

Such methods have of course always been used in the theory of gases. But the manner in which the jumble of molecules moving in all directions was replaced by groups of average constitution, the way in which one tried to explain the distribution of velocities by means of *Maxwell's* law and to apply the calculus of probabilities to the mean number and to the effect of the collisions was not always unobjectionable. Moreover, all these methods referred for the most part to the special case, exclusive to the theory of gases, of a sys-

ten Teilchen und waren somit auf die Theorie der Flüssigkeiten und festen Körper nicht ohne weiteres anwendbar. Es besteht also gewiß ein Bedürfnis, die Grundlagen der statistischen Mechanik unabhängig von ihrem Anwendungsgebiete auf streng mathematischer Grundlage zu entwickeln, und dies ist eben die Aufgabe, die der jüngst verstorbene große amerikanische Physiker sich in seinem hier vorliegenden letzten Werke gestellt hat, um hiermit der kinetischen Theorie einen ähnlichen Dienst zu leisten, wie er ihn in seinen grundlegenden „Thermodynamischen Studien“ der phänomenologischen Thermochemie erwiesen hat. Das behandelte Problem bietet in neuester Zeit noch ein erhöhtes Interesse: seitdem man begonnen hat, auch die Elektrizität atomistisch aufzufassen, wird man auch in der Theorie der bewegten „Elektronen“ statistische Methoden nicht entbehren können.

Ebenso wenig wie *Maxwell*<sup>1</sup> bei der Untersuchung der intramolekularen Bewegung legt *Gibbs* seiner Theorie eine speziellere Molekularvorstellung zugrunde, sondern ausschließlich die allgemeine Form der dynamischen Gleichungen, welche für ein System von  $n$  Freiheitsgraden unter der Einwirkung von inneren Potentialkräften in der *Hamiltonschen* Form geschrieben werden können:

$$\dot{q}_i = \frac{dq_i}{dt} = \frac{\partial \varepsilon}{\partial p_i}, \quad \dot{p}_i = \frac{dp_i}{dt} = -\frac{\partial \varepsilon}{\partial q_i}, \quad (6)$$

235 wo  $q_1, q_2 \dots q_n$  die Koordinaten,  $p_1, p_2 \dots p_n$  die Impulse (momenta) und  $\varepsilon$  die Gesamtenergie des Systemes bedeutet. Die Koordinaten für sich | bestimmen die Konfiguration, die Impulse den Geschwindigkeitszustand, beide zusammen die „Phase“ des Systems für einen variablen Zeitpunkt  $t$ . Durch die Anfangsphase zur Zeit  $t = t_0$  ist nun die Phase zu jeder anderen Zeit  $t$  und damit die ganze Bewegung des Systems vollständig bestimmt. Läßt man also die Anfangsphase variieren innerhalb eines gewissen  $2n$ -dimensionalen Gebietes  $g_0$ , so variiert auch die Phase zur Zeit  $t_1$  innerhalb eines entsprechenden Gebietes  $g_1$ , und nach einem bekannten Satze von *Liouville* hat das  $2n$  fache Integral  $\int \dots \int dp_1 \dots dp_n dq_1 \dots dq_n$ , das von der Wahl des Koordinatensystemes unabhängig ist und als „Phasenausdehnung“ (extension-in-phase) bezeichnet wird, für beide Gebiete  $g_0$  und  $g_1$  denselben Wert, d. h. es ist für bewegte Systeme in der Zeit konstant. Dieser Satz wird bezeichnet als der „Satz von der Erhaltung der Phasenausdehnung“ und ist für alles Folgende von grundlegender Bedeutung. Anstatt eines einfachen Systemes denken wir uns eine große Zahl von Systemen mit verschiedenen Anfangsphasen in der Weise verteilt, daß der Quotient aus der Anzahl der Systeme, die sich innerhalb eines Gebietes  $g$  befinden, durch die Ausdehnung dieses Gebietes bei entsprechender Verkleinerung einer Grenze  $D$  zustrebt, die im allgemeinen eine Funktion der Phase  $(p, q)$  ist und als „Phasendichte“ (density-in-phase) bezeichnet wird. Der Quotient  $P$  dieser Phasendichte durch die Anzahl  $N$  aller überhaupt betrachteten Systeme heißt der „Wahrscheinlich-

<sup>1</sup> *Maxwell*, *Cambr. Trans.* 12 (1879), p. 547; *Papers* 2, p. 713.

tem of particles in linear motion, and hence were not applicable to the theory of liquids and solid bodies without further ado. Thus there certainly exists a need for developing the foundations of statistical mechanics independently of its domain of application on the basis of a rigorous mathematical foundation, which is just the task the great, recently deceased American physicist has set himself in his last work under review here, in which he attempted to do for the kinetic theory what he had done for the phenomenological thermochemistry in his seminal *1892*. The problem under consideration has drawn increased attention recently: with the emergence of an atomistic conception of electricity, it will be impossible to do without statistical methods also in the theory of “electrons” in motion.

Like *Maxwell*<sup>1</sup> in the investigation of intramolecular motion, *Gibbs* does not base his theory upon a specific conception of molecules but only upon the general form of the dynamical equations, which, for a system of  $n$  degrees of freedom under the action of internal potential forces, can be written in *Hamiltonian* form:

$$\dot{q}_i = \frac{dq_i}{dt} = \frac{\partial \varepsilon}{\partial p_i}, \quad \dot{p}_i = \frac{dp_i}{dt} = -\frac{\partial \varepsilon}{\partial q_i}, \quad (6)$$

where  $q_1, q_2 \dots q_n$  denote the coordinates,  $p_1, p_2 \dots p_n$  the momenta, and  $\varepsilon$  the total energy of the system. The coordinates themselves determine the configuration, and the momenta determine the velocity state, and both of them together determine the “phase” of the system for a variable point in time  $t$ . The initial phase at time  $t = t_0$  now completely determines the phase at any other time  $t$ , and hence the entire motion of the system. Thus, if we let the initial phase vary within a certain  $2n$ -dimensional region  $g_0$ , then the phase at time  $t_1$  also varies within a corresponding region  $g_1$ , and, by a well-known theorem of *Liouville*, the  $2n$ -fold integral  $\int \dots \int dp_1 \dots dp_n dq_1 \dots dq_n$ , which is independent of the choice of the coordinate system and is called “extension-in-phase”, has the same value for both regions  $g_0$  and  $g_1$ , i. e., it is constant over time for systems in motion. This theorem is called the “theorem of the preservation of the extension-in-phase” and is of basic significance for all that follows. Instead of a simple system, let us envision a great number of systems with different initial phases distributed so that the quotient from the number of systems lying within a region  $g$ , and the expansion of this region upon corresponding reduction, converges to a limit  $D$ , which, in general, is a function of the phase  $(p, q)$  and is called the “density-in-phase”. The quotient  $P$  of this density-in-phase by the number  $N$  of all systems under

<sup>1</sup> *Maxwell 1879; Maxwell 1890*, pp. 713 ff.

keitskoeffizient“ (coefficient of probability), dessen natürlicher Logarithmus  $\eta$  ist der „Wahrscheinlichkeitsexponent“ (index of probability), und der Ausdruck  $P dp_1 \dots dq_n = e^\eta dp_1 \dots dq_n$  stellt den Wert der Wahrscheinlichkeit dar, daß ein gegebenes System sich zwischen den durch das Differentialprodukt angedeuteten Phasengrenzen befindet. Denken wir uns nun ein beliebiges System gleichzeitig mit den anderen Systemen, die von benachbarten Phasen ausgehen, gemäß den Bewegungsgleichungen verändert, so muß dabei mit der Phasenausdehnung und der Anzahl der Systeme auch die Phasendichte  $D$  und daher auch  $P$  und  $\eta$  ungeändert bleiben. Ist also die Phasendichte zu einer beliebigen Zeit  $t$  als Funktion der Phase  $(p, q)$  gegeben, so ist für die betrachtete „Gesamtheit“ (ensemble) die „Phasenverteilung“ (distribution-in-phase) für jede andere Zeit  $t$  bestimmt vermöge der partiellen Differentialgleichung

$$\left(\frac{dD}{dt}\right)_{p,q} + \sum \left(\frac{\partial D}{\partial p_i} \dot{p}_i + \frac{\partial D}{\partial q_i} \dot{q}_i\right) = 0 \quad (21)$$

oder

$$\frac{\partial D}{\partial t} = \sum \left(\frac{\partial D}{\partial p_i} \frac{\partial \varepsilon}{\partial q_i} - \frac{\partial D}{\partial q_i} \frac{\partial \varepsilon}{\partial p_i}\right),$$

die als die „Grundgleichung der statistischen Mechanik“ bezeichnet werden kann. Hier bedeutet  $\frac{\partial D}{\partial t}$  oder genauer  $\left(\frac{dD}{dt}\right)_{p,q}$  die zeitliche Veränderung von  $D$  für konstante  $p, q$ , d. h. an einer bestimmten Stelle des Phasenraumes.

Die Phasendichte bleibt an jeder Stelle  $(p, q)$  ungeändert:  $\left(\frac{dD}{dt}\right)_{p,q} = 0$ , wenn die rechte Seite der letzten Gleichung verschwindet, und in diesem Falle sagen wir, unsere Gesamtheit sei in „statistischem Gleichgewicht“ (statistical equilibrium). Dies wird insbesondere dann eintreten, wenn die Phasendichte konstant oder eine Funktion der Gesamt- | energie ist. Ob eine Gesamtheit in statistischem Gleichgewichte ist oder nicht, hängt demnach nicht von der Beschaffenheit der einzelnen Systeme selbst ab, sondern lediglich von der willkürlich angenommenen anfänglichen Phasenverteilung, d. h. von der Anzahl der Systeme, die wir von jedem Phasenelement aus ihre Bewegung beginnen lassen. Die in einem einzelnen Systeme stattfindenden Vorgänge sind daher ganz unabhängig davon, ob das betrachtete System einer in statistischem Gleichgewichte stehenden Gesamtheit angehört oder nicht. Zur Erleichterung der meisten Untersuchungen wird es zweckmäßig sein, Verteilungen in statistischem Gleichgewichte vorzugsweise zu betrachten, und es bietet sich eine weitere Vereinfachung, wenn wir die Phasendichte kontinuierlich nach 0 abnehmen lassen, während die Koordinaten und Impulse  $q, p$  ins Unendliche wachsen. Beiden Forderungen genügt eine Verteilung, in welcher der Wahrscheinlichkeitsexponent eine lineare Funktion der Energie ist:

$$\eta = \log P = \frac{\psi - \varepsilon}{\theta} \quad \text{oder} \quad P = e^\eta = e^{\frac{\psi - \varepsilon}{\theta}},$$

consideration is called the “coefficient of probability”, whose natural logarithm  $\eta$  is the “index of probability”, and the expression  $P dp_1 \dots dq_n = e^\eta dp_1 \dots dq_n$  represents the value of the probability that a given system lies between the phase boundaries indicated by the differential product. Let us now consider any system together with the other systems starting from neighboring phases, all modified in accordance with the equations of motion. Then along with the extension-in-phase and the number of the systems the density-in-phase  $D$ , too, and hence also  $P$  and  $\eta$ , must remain unchanged. Thus, if the density-in-phase is given at any time  $t$  as function of the phase  $(p, q)$ , then, for any other time  $t$ , the “distribution-in-phase” is determined for the considered “ensemble” by virtue of the partial differential equation

$$\left(\frac{dD}{dt}\right)_{p,q} + \sum \left(\frac{\partial D}{\partial p_i} \dot{p}_i + \frac{\partial D}{\partial q_i} \dot{q}_i\right) = 0 \quad (21)$$

or

$$\frac{\partial D}{\partial t} = \sum \left(\frac{\partial D}{\partial p_i} \frac{\partial \varepsilon}{\partial q_i} - \frac{\partial D}{\partial q_i} \frac{\partial \varepsilon}{\partial p_i}\right),$$

which can be called the “basic equation of statistical mechanics”. Here,  $\frac{\partial D}{\partial t}$ , or more precisely,  $\left(\frac{dD}{dt}\right)_{p,q}$ , denotes the temporal change of  $D$  for constant  $p, q$ , i. e., at a particular point of the phase space. The density-in-phase remains unchanged at every point  $(p, q)$ :  $\left(\frac{dD}{dt}\right)_{p,q} = 0$ , if the right side of the last equation vanishes, and in this case we say that our ensemble is in “statistical equilibrium”. This will occur in particular when the density-in-phase is constant or a function of the total energy. Thus, whether or not an ensemble is in statistical equilibrium does not depend on the constitution of the individual systems themselves but only on the arbitrarily assumed initial distribution-in-phase, i. e., on the number of the systems whose motions we let begin from every phase element. Therefore, the processes occurring in an individual system are entirely independent of whether or not the considered system belongs to an ensemble in statistical equilibrium. To simplify most investigations, it will be convenient to preferably consider distributions in statistical equilibrium, and a further simplification is afforded by letting the density-in-phase continually decrease toward 0 while the coordinates and momenta  $q, p$  increase to infinity. Both requirements are satisfied by a distribution in which the index of probability<sup>2</sup> is a linear function of the energy:

$$\eta = \log P = \frac{\psi - \varepsilon}{\theta} \quad \text{or} \quad P = e^\eta = e^{\frac{\psi - \varepsilon}{\theta}},$$

<sup>2</sup> [Zermelo’s term “Wahrscheinlichkeitsexponent” is translated here and in the following by the Gibbsian term “index of probability”.]



wo  $\psi$  und  $\theta$  Konstanten sind und  $\psi$  als Funktion von  $\theta$  so zu bestimmen ist, daß die Gesamtwahrscheinlichkeit, d. h. das Integral

$$\int e^{\frac{\psi - \varepsilon}{\theta}} dp_1 \dots dq_n,$$

über alle möglichen Phasen bis ins Unendliche erstreckt, den Wert 1 hat. Eine solche Verteilung bezeichnet *Gibbs* als eine „kanonische“ und beschränkt seine weiteren Untersuchungen zumeist auf solche Gesamtheiten. Die Analogie einer kanonischen Verteilung mit dem *Maxwellschen* Verteilungsgesetze in der Gastheorie ist unmittelbar einleuchtend, nur handelt es sich hier nicht um verschiedene Teilchen desselben Systems, sondern um lauter voneinander unabhängige Systeme. Speziell für solche kanonischen Verteilungen werden nun „Durchschnittswerte“, d. h. Integrale der Form

$$\bar{u} = \int \dots \int u e^{\frac{\psi - \varepsilon}{\theta}} dp_1 \dots dq_n$$

für verschiedene Phasenfunktionen  $u$  sowie ihre „mittleren Fehler“ d. h. ihre durchschnittlichen Abweichungen vom Durchschnittswerte berechnet.

Die für eine kanonische Verteilung charakteristische Konstante, den „Verteilungsmodul“  $\theta$ , betrachtet *Gibbs* in seiner Analogie mit der Temperatur eines thermischen Systems. Werden nämlich zwei kanonische Gesamtheiten zu einer dritten Gesamtheit in der Weise vereinigt, daß der Wahrscheinlichkeitskoeffizient des Gesamtsystems durch das Produkt  $P_1 P_2$  der beiden gegebenen Koeffizienten für dieselbe Phase dargestellt wird, so ist die entstehende Gesamtheit wieder eine kanonische, also in statistischem Gleichgewichte, wenn die Verteilungsmoduln  $\theta_1$  und  $\theta_2$  der beiden Gesamtheiten gleich sind, ebenso wie zwei thermische Systeme nur dann miteinander im Gleichgewichte sein können, wenn ihre Temperaturen gleich sind. Denkt man sich ferner den Verteilungsmodul  $\theta$  einer kanonischen Gesamtheit variiert, betrachtet also eine kontinuierliche Schar solcher Gesamtheiten mit verschiedenen Moduln, so ändert sich auch  $\psi$  und ergibt für die Durchschnittswerte  $\bar{\varepsilon}$  und  $\bar{\eta}$  der Energie und des Wahrscheinlichkeitsexponenten die Beziehung

$$d\psi = \frac{\psi - \bar{\varepsilon}}{\theta} d\theta = \bar{\eta} d\theta.$$

237 | Werden gleichzeitig auch die „äußeren Koordinaten“ verändert, d. h. die Parameter  $a_1, a_2, a_3, \dots$ , welche in dem Ausdrucke  $\varepsilon$  der Gesamtenergie außer den Koordinaten und Impulsen  $p, q$  als Veränderliche auftreten, sodaß sie für alle Systeme einer einzigen Gesamtheit dieselben Werte, für jede veränderte Gesamtheit aber andere Werte haben, so kommen noch weitere Glieder hinzu, welche die Durchschnittswerte der „äußeren Kräfte“  $A_1 = -\frac{\partial \varepsilon}{\partial a_1}$ ,  $A_2 = -\frac{\partial \varepsilon}{\partial a_2}, \dots$  darstellen und mit  $\bar{A}_1, \bar{A}_2, \dots$  bezeichnet werden mögen, und

where  $\psi$  and  $\theta$  are constants and  $\psi$  is to be determined as a function of  $\theta$  so that the total probability, i. e., the integral

$$\int e^{\frac{\psi-\varepsilon}{\theta}} dp_1 \dots dq_n,$$

extended over all possible phases to infinity, has the value 1. *Gibbs* refers to a distribution of this sort as “canonical” and restricts his further investigations for the most part to such ensembles. The analogy of a canonical distribution with *Maxwell’s* law of distribution in the theory of gases is immediately evident, but it is concerned here with mutually independent systems rather than different particles of the same system. Now, specifically for such canonical distributions “mean values” are calculated, i. e., integrals of the form

$$\bar{u} = \int \dots \int u e^{\frac{\psi-\varepsilon}{\theta}} dp_1 \dots dq_n$$

for different phase functions  $u$  and their “mean errors”, i. e., their average deviations from the mean value.

*Gibbs* considers the constant characteristic of a canonical distribution, the “distribution module”  $\theta$ , in his analogy with the temperature of a thermal system. For if two canonical ensembles are united with a third ensemble so that the coefficient of probability of the total system is represented by the product  $P_1 P_2$  of the two given coefficients for the same phase, then the ensemble obtained is in turn a canonical ensemble, and hence in statistical equilibrium, if the distribution modules  $\theta_1$  and  $\theta_2$  of the two ensembles are identical, just like two thermal systems can be in equilibrium with one another only if their temperatures are identical. Furthermore, if we assume that the distribution module  $\theta$  of a canonical ensemble varies, and hence if we consider a continuous family of such ensembles with different modules, then  $\psi$ , too, changes and, for the mean values  $\bar{\varepsilon}$  and  $\bar{\eta}$  of the energy and of the index of probability, we get the relation

$$d\psi = \frac{\psi - \bar{\varepsilon}}{\theta} d\theta = \bar{\eta} d\theta.$$

If we also alter the “external coordinates”, i. e., the parameters  $a_1, a_2, a_3 \dots$  occurring in the expression  $\varepsilon$  of the total energy besides the coordinates and momenta  $p, q$  as variables, so that they have the same values for all systems of a single ensemble but different values for each altered ensemble, then further terms are added which represent the mean values of the “external forces”  $A_1 = -\frac{\partial \varepsilon}{\partial a_1}$ ,  $A_2 = -\frac{\partial \varepsilon}{\partial a_2}$ ,  $\dots$  and which we denote by  $\bar{A}_1, \bar{A}_2, \dots$ , and we

wir erhalten die Formeln

$$d\psi = \bar{\eta}d\theta - \bar{A}_1 da_1 - \bar{A}_2 da_2 - \dots$$

oder wegen  $d\psi - d\bar{\varepsilon} = d(\theta\bar{\eta})$  schließlich

$$d\bar{\varepsilon} = -\theta d\bar{\eta} - \bar{A}_1 da_1 - \bar{A}_2 da_2 - \dots \quad (114)$$

ganz analog der Differentialbeziehung des zweiten Hauptsatzes, wenn  $\theta$  der Temperatur,  $\bar{\varepsilon}$  der Energie und  $\bar{\eta}$  der negativen Entropie entspricht. Diese Analogie bleibt aber rein äußerlich und kann uns keine Aufklärung über den mechanischen Charakter der thermodynamischen Gleichung bieten. Denn die Größe  $\bar{\eta}$ , der Durchschnittswert des Wahrscheinlichkeitsexponenten, ist gar keine Eigenschaft der einzelnen mechanischen Systeme, auch keine durchschnittliche Eigenschaft, welche das allen Gemeinsame oder Vorwiegende zum Ausdruck bringen könnte, sondern sie hängt wie der Wahrscheinlichkeitsexponent selbst ausschließlich von der willkürlichen anfänglichen Phasenverteilung ab, während die Entropie eines Systems doch immer durch die Natur dieses einzelnen Systems selbst bestimmt sein muß. Ebenso wenig kann das Glied  $\theta d\bar{\eta}$  im eigentlichen Sinne die übergehende Wärme repräsentieren, da es sich lediglich auf eine Veränderung der Phasenverteilung, auf den Übergang zu einer neuen Gesamtheit, gar nicht auf eine Wechselwirkung der Systeme mit der Außenwelt bezieht. Will man in einwandfreier Weise die Eigenschaften eines mechanischen Systems statistisch untersuchen, so muß man immer die Gesamtheit beibehalten und nur innerhalb dieser Gesamtheit die Änderungen der Durchschnittswerte bestimmen. Denn das einzelne System in seinem durchschnittlichen oder wahrscheinlichen Verhalten, nicht eine zwischen solchen Systemen willkürlich festgesetzte Beziehung sollte doch das mechanische Analogon eines physikalisch veränderten Naturkörpers darstellen.

Im 10. Kapitel untersucht der Verfasser eine andere Analogie mit Hilfe der „mikrokanonischen Phasenverteilung“, d. h. einer Phasenverteilung, in welcher alle Systeme die gleiche Energie haben, und die auch als ein Grenzfall einer kanonischen Gesamtheit aufgefaßt werden kann, indem man den Ausdruck  $\eta = c - \frac{(\varepsilon - \varepsilon')^2}{\omega^2}$  für sehr große Werte von  $\omega$  als Wahrscheinlichkeitsexponenten ansetzt. Jedem Energiewerte  $\varepsilon$  entspricht nun ein bestimmter Wert des im 8. Kapitel eingeführten Integrales

$$V = \int \dots \int dp_1 \dots dq_n$$

erstreckt über alle Phasen mit einer Energie  $< \varepsilon$ . Somit ist  $V$ , die Phasenausdehnung unter einer Energiegrenze  $\varepsilon$ , sowie auch der abgeleitete Wert  $\varphi = \log \frac{\partial V}{\partial \varepsilon}$  eine Funktion von  $\varepsilon$  und den äußeren Koordinaten  $a_1, a_2, \dots$ ,

238 | also konstant in jeder mikrokanonischen Gesamtheit. Bezeichnet man nun

obtain the formulas

$$d\psi = \bar{\eta}d\theta - \bar{A}_1 da_1 - \bar{A}_2 da_2 - \dots$$

or eventually, on account of  $d\psi d\bar{\varepsilon} = d(\theta\bar{\eta})$ ,

$$d\bar{\varepsilon} = -\theta d\bar{\eta} - \bar{A}_1 da_1 - \bar{A}_2 da_2 - \dots \quad (114)$$

in a manner entirely analogous to the differential relation of the second law, where  $\theta$  corresponds to the temperature,  $\bar{\varepsilon}$  to the energy and  $\bar{\eta}$  to the negative entropy. But this analogy remains entirely superficial and cannot clarify the mechanical character of the thermodynamical equation. For the magnitude  $\bar{\eta}$ , the mean value of the index of probability, is not at all a property of the individual mechanical systems. Nor is it an average property capable of expressing what is common to or dominant in all. Rather, like the index of probability itself, it depends exclusively on the arbitrary initial distribution-in-phase, while the entropy of a system must always be determined by the nature of this individual system itself. Likewise, the term  $\theta d\bar{\eta}$  cannot actually represent the transmitted heat, since it only refers to a change of the distribution-in-phase, to the transition to a new ensemble, and not at all to an interaction of the systems with the external world. Anyone wishing to statistically investigate in an unobjectionable manner the properties of a mechanical system must always retain the ensemble and only determine the changes of the mean values within this ensemble. For, after all, it is the individual system in its average or probable behavior, not a relation arbitrarily stipulated among such systems, that should be the mechanical analogy of a physically altered natural body.

In chapter 10, the author examines another analogy by means of the “microcanonical distribution-in-phase”, i. e., a distribution-in-phase in which all systems have the same energy and which can also be regarded as a borderline case of a canonical ensemble by setting the expression  $\eta = c - \frac{(\varepsilon - \varepsilon')^2}{\omega^2}$  as index of probability for very large values of  $\omega$ . To every energy value  $\varepsilon$  there now corresponds a particular value of the integral introduced in chapter 8

$$V = \int \dots \int dp_1 \dots dq_n$$

extended over all phases with energy  $< \varepsilon$ . Thus,  $V$ , the extension-in-phase below some energy limit  $\varepsilon$ , as well as the derived value  $\varphi = \log \frac{\partial V}{\partial \varepsilon}$ , is a function of  $\varepsilon$  and the external coordinates  $a_1, a_2, \dots$ , and hence constant in every

mit  $\overline{A_1}_\varepsilon$ ,  $\overline{A_2}_\varepsilon$  usw. die in der mikrokanonischen Gesamtheit genommenen Durchschnittswerte der äußeren Kräfte, so gelangt man zu der Gleichung

$$d\varepsilon = e^{-\varphi} V d \log V - \overline{A_1}_\varepsilon da_1 - \overline{A_2}_\varepsilon da_2 - \dots, \quad (418)$$

welche wieder der Gleichung des zweiten thermodynamischen Hauptsatzes entspricht, wenn  $e^{-\varphi} V = V / \frac{\partial V}{\partial \varepsilon}$  die Temperatur und  $\log V$  die Entropie repräsentiert. Diese zweite Analogie entspricht nun schon sehr viel besser einem wirklichen Naturvorgange, da alle hier vorkommenden Größen wirkliche Eigenschaften der einzelnen Systeme sind, bestimmt durch den Ausdruck, welcher die Energie als Funktion der äußeren und inneren Koordinaten und der Impulse darstellt. Bezeichnet man ferner mit  $\varepsilon_p$  die kinetische Energie des Systems, so gilt für ihren Durchschnittswert in der mikrokanonischen Gesamtheit die Beziehung

$$\frac{2}{n} \overline{\varepsilon_p}_\varepsilon = e^{-\varphi} V, \quad (377)$$

es wird also wie in der kinetischen Gastheorie die lebendige Kraft der bewegten Teilchen ein Analogon der Temperatur. Dagegen gilt für solche mikrokanonischen Gesamtheiten nicht allgemein das Analogon des Satzes, daß zwei Körper von gleicher Temperatur vereinigt wieder ein System von derselben Temperatur ergeben.

Einen wesentlichen Bestandteil der Theorie bildet das 12. Kapitel, in welchem die Bewegung der Systeme und Gesamtheiten in langen Zeiten untersucht wird. Aus dem Gesetze von der Erhaltung der Phasenausdehnung wird der Satz von *Poincaré* hergeleitet, daß ein innerhalb endlicher Grenzen eingeschlossenes und sich selbst überlassenes, nur den *Hamiltonschen* Gleichungen folgendes mechanisches System im *Poissonschen* Sinne *stabil* ist. Betrachten wir nämlich ein beliebig kleines Gebiet  $g$  von Anfangsphasen von endlicher Phasenausdehnung, so wird wenigstens ein Teil, sogar der größte Teil dieser Phasen nach Ablauf einer beliebig langen Zeit  $T$ , wenn auch nicht genau in die Anfangsphase, so doch in dasselbe Gebiet  $g$  zurückkehren, die Bewegung wird also für den überwiegenden Teil der Phasen einen quasiperiodischen Charakter haben. Somit steht nicht zu erwarten, daß sich bei Bewegungen dieser Art eine Tendenz nach irgend einem Endzustande geltend machen werde. Gleichwohl versucht der Verfasser ein solches Verhalten wenigstens wahrscheinlich, plausibel zu machen, indem er seinen Gesamtheiten, sofern sie nicht schon in statistischem Gleichgewichte sind, eine Tendenz nach einem Endzustande statistischen Gleichgewichtes zuschreiben möchte. Für den Fall des statistischen Gleichgewichtes ist nun, wie im elften Kapitel gezeigt wird, der durchschnittliche Wahrscheinlichkeitsexponent  $\overline{\eta}$  *kleiner* als bei jeder anderen Verteilung. Es müßte sich also der Wert  $\overline{\eta}$  mit unbegrenzt wachsender Zeit seiner unteren Grenze  $\overline{\eta}_0$  beliebig nähern. Dem gegenüber steht aber die Tatsache, daß der Wahrscheinlichkeitsexponent, welcher der jeweiligen Phase eines bewegten Systemes entspricht, und damit auch der mittlere Wahr-

microcanonical ensemble. If we now denote the mean values of the external forces taken in the microcanonical ensemble by  $\overline{A_1}]_\varepsilon$ ,  $\overline{A_2}]_\varepsilon$  etc., then we obtain the equation

$$d\varepsilon = e^{-\varphi} V d \log V - \overline{A_1}]_\varepsilon da_1 - \overline{A_2}]_\varepsilon da_2 - \dots, \quad (418)$$

which in turn corresponds to the equation of the second law of thermodynamics, assuming that  $e^{-\varphi} V = V / \frac{\partial V}{\partial \varepsilon}$  represents the temperature, and  $\log V$  the entropy. Now, this second analogy is already in much better accord with a real process in nature, since all magnitudes occurring here are real properties of the individual systems, determined by the expression representing the energy as a function of the external and internal coordinates and of the momenta. If, furthermore, we denote the kinetic energy of the system by  $\varepsilon_p$ , then for its mean value in the microcanonical ensemble the following relation holds:

$$\frac{2}{n} \overline{\varepsilon_p}]_\varepsilon = e^{-\varphi} V, \quad (377)$$

and hence, just as in the kinetic theory of gases, the living force of the particles in motion turns into an analogue of temperature. By contrast, an analogue to the theorem that uniting two bodies of the same temperature results in a system of the same temperature does not hold generally for such microcanonical ensembles.

An essential part of the theory is formed by chapter 12, which examines the motion of the systems and ensembles over extended periods of time. From the law of the preservation of the extension-in-phase *Gibbs* deduces *Poincaré's* theorem that a mechanical system enclosed within finite boundaries, when left to its own devices, only following the *Hamiltonian* equations, is *stable* in *Poisson's* sense of the word. For if we consider an arbitrarily small region  $g$  of initial phases of finite extension-in-phase, then, after an arbitrarily long period of time  $T$  has passed, at least a few, and even the majority, of these phases will return, if not exactly to the initial phase, at least into the same region  $g$ . Hence, the motion will have quasi-periodic character for the majority of phases. It is therefore not to be expected that a tendency towards any final state will assert itself for motions of this kind. Nonetheless, the author tries to render such behavior at least probable or plausible, by wishing to assign to his ensembles a tendency toward a final state of statistical equilibrium, provided they are not already in statistical equilibrium. As is shown in chapter 11, the mean index of probability  $\overline{\eta}$  is *smaller* for the case of the statistical equilibrium than for any other distribution. Hence, the value  $\overline{\eta}$  would have to get arbitrarily close to its lower boundary  $\overline{\eta}_0$  at indefinitely increasing time. But this runs counter to the fact that the index of probability, which corresponds to the respective phase of a system in motion, and hence also the

239

scheinlichkeitsexponent der ganzen Gesamtheit auf Grund des Gesetzes von der Erhaltung der Phasendichte von der Zeit unabhängig, eine Konstante ist. Hier sucht sich nun der Verfasser durch Umdeutung und Analogie zu helfen. Er denkt sich in einem geschlossenen Raum zwei nicht-diffundierbare inkompressible Flüssigkeiten von gleicher hydrodynamischer Beschaffenheit, eine gefärbte und eine ungefärbte, durcheinander gerührt. Dann wird nach einiger Zeit, im allgemeinen wenigstens, eine nahezu homogene Mischung entstehen, obwohl jedes Flüssigkeitsteilchen nach wie vor entweder der einen oder der anderen Flüssigkeit angehören muß, die Dichtigkeit des Farbstoffes also in jedem kleinsten Teilchen dieselbe geblieben ist. Das mittlere Quadrat dieser Dichte, das im Falle wahrer Homogenität seinen Minimalwert annehmen müßte, wird also dasselbe sein wie am Anfang, wenn die Dichte in jedem Augenblicke durch unbegrenzte Verkleinerung des Volumenelementes bestimmt wird. Berechnet man dagegen die mittlere Dichte für endliche wenn auch beliebig kleine Volumenelemente, die nach einer gewissen Zeit eine große Menge verschiedenartiger Flüssigkeitsteilchen enthalten werden, so erscheint die so bestimmte Dichte im Endzustande nahezu homogen, und ihr mittleres Quadrat hat gegen den Anfangszustand abgenommen. In analoger Weise glaubt der Verfasser durch Vertauschung der beiden Grenzübergänge der Folgerung von der Konstanz des mittleren Wahrscheinlichkeitsexponenten entgehen zu können, indem er nicht ohne Berechtigung der betrachteten Bewegung in dem Sinne „Instabilität“ zuschreibt, daß im allgemeinen anfangs benachbarte Phasen sich im Laufe der Zeit immer weiter voneinander entfernen und anfangs getrennte sich nähern werden. Diese Analogie ist aber irreführend. In dem vorliegenden Beispiele handelt es sich wesentlich um einen Gegensatz zwischen kontinuierlicher und diskontinuierlicher Dichtigkeitsverteilung. Eine anfangs vorhandene Diskontinuitätsfläche von beschränkter Ausdehnung wird durch fortgesetzte Bewegung in eine schwammartig verwickelte Fläche von ungeheurer Ausdehnung auseinander gezogen und über den ganzen Raum verbreitet, sodaß jeder nicht allzu kleine Raumteil schroffe Übergänge von der einen in die andere Flüssigkeit in sich enthält, seine mittlere Dichtigkeit aber von der des benachbarten nur unbedeutend abweicht. Es wird also durch das Umrühren die Inhomogenität aus dem Großen ins Kleine übertragen, während ihre durch das mittlere Quadrat der Farbstoffdichte gemessene Gesamtgröße ungeändert bleibt. Von einem solchen Gegensatz zwischen kontinuierlichen und diskontinuierlichen Dichtigkeitsänderungen ist aber bei den von *Gibbs* betrachteten Gesamtheiten durchaus nicht die Rede. Vielmehr wird die Phasendichte von vornherein als eine stetige Funktion des Ortes im Phasenraume aufgefaßt, und es liegt kein Grund vor, anzunehmen, daß sie in den kleinsten Teilchen immer inhomogener werden sollte, um in den größeren Teilen im Mittel homogener werden zu können. Daß sich der Verfasser jede Phasenausdehnung bei der ursprünglichen Einführung der Phasendichte nicht stetig sondern durch eine endliche wenn auch sehr große Zahl von Systemen erfüllt denkt, tut nichts zur Sache, da dies anfangs so gut wie später gelten müßte, seine Formeln wenigstens haben es nur mit einer im allgemeinen stetigen

mean index of probability of the entire ensemble by virtue of the law of the preservation of the density-in-phase, is independent of time, is a constant. The author tries to find a way here by reinterpretation and analogy. He envisions two non-diffusible incompressible liquids of identical hydrodynamical constitution, one dyed and one undyed, blended in a closed space. Then, after some time, at least in general, an almost homogeneous mixture is formed, even though each liquid-particle must belong to either one or the other liquid, and hence the density of the colorant in every smallest particle has remained the same. Thus, the mean square of this density, which would have to assume its minimal value in the case of true homogeneity, will be the same as in the beginning, if the density is determined at every instant by unlimited decrease of the volume element. If, on the other hand, we calculate the mean density for finite, if arbitrarily small, volume elements that will, after a certain time, contain a great amount of various liquid-particles, then the density thus determined appears almost homogeneous in the final state, and its mean square has decreased toward the initial state. The author believes that it is possible to analogously avoid the conclusion about the constancy of the mean index of probability through a switch of the two passages to the limit by assigning, not without justification, "instability" to the considered motion so that, in general, initially neighboring phases will move ever further apart from one another over the course of time, and initially distant phases will approach one another. But this analogy is misleading. The present example hinges on a contrast between continuous and discontinuous density distributions. An initially given discontinuity surface of limited extension is stretched into a sponge-like entangled surface of immense extension by continued motion and is spread over the entire space so that each part of the space that is not too small contains jagged transitions from one liquid to the other while its mean density deviates from that of the neighboring part only insignificantly. Thus, the non-homogeneity is carried over from the large to the small scale by the stirring, while its total magnitude, which is measured by the mean square of the colorant density, remains unaltered. But, when considering his ensembles, *Gibbs* does not at all refer to such a contrast between continuous and discontinuous changes in density. Rather, the density-in-phase is conceived of at the outset as a continuous function of the position in the phase space, and there is no reason to assume that it should become ever more inhomogeneous in the smallest particles in order to be able to become on average more homogeneous in the greater parts. That the author assumes that each extension-in-phase is not continuous at the original introduction of the density-in-phase but satisfied by a finite if very large number of systems is irrelevant here, for this should hold true at the beginning just as at a later stage. At least his formulas only refer to a phase distribution, which, in



Phasenverteilung zu tun. Aus der oben zitierten Grundgleichung (21) der statistischen Mechanik folgt mathematisch streng, daß auch  $\eta = \log \frac{D}{N}$  der Differentialgleichung genügt

$$\frac{\partial \eta}{\partial t} = \sum \left( \frac{\partial \eta}{\partial p_i} \frac{\partial \varepsilon}{\partial q_i} - \frac{\partial \eta}{\partial q_i} \frac{\partial \varepsilon}{\partial p_i} \right)$$

240 | und daher  $\int \dots \int \eta dp_1 \dots dq_n$  eine Bewegungsinvariante ist, d. h.  $\bar{\eta} = \text{const.}$ , und dies gilt ganz unabhängig vom Koordinatensystem, mag man die Volumenelemente ursprünglich größer oder kleiner wählen, die Phasendichte durch den einen oder den anderen Grenzübergang bestimmen.

Wäre es aber auch gelungen, durch geeignete Umdeutung des Dichtigkeitsbegriffes das gewünschte Resultat zu erzielen, so stände es doch jedenfalls im Widerspruche mit den anfangs gegebenen Grundgleichungen, d. h. mit der gesamten vorher entwickelten Theorie. Man müßte geradezu von vorn anfangen und eine neue Theorie begründen, in welcher der Satz von der Herstellung des statistischen Gleichgewichtes gelten könnte. Ganz unstatthaft ist es aber, wie der Verfasser es im folgenden Kapitel tut, Formeln der alten und der neuen Theorie zu kombinieren, indem er z. B. in (462) die Annahme  $\bar{\eta}''_{12} \leq \bar{\eta}'_{12}$  zu den Folgerungen aus früheren Entwicklungen hinzufügt. Mit widersprechenden Voraussetzungen kann man freilich alles beweisen, was man will, beweist aber damit in Wirklichkeit nur die Unzulässigkeit eines solchen Verfahrens. Aber noch ein Weiteres hätte ich gegen die Formel (462) einzuwenden. Angenommen selbst, es wäre bewiesen, daß  $\bar{\eta}$  mit der Zeit seinem Minimalwerte zustrebe, so müßte dies, wie der Verfasser selbst zugibt (*Gibbs*, p. 150, deutsche Bearbeitung S. 153) für abnehmende Zeiten ebenso gut gelten wie für wachsende, und die Grenzwerte für  $t = +\infty$  und für  $t = -\infty$  wären dieselben. Dann ist es aber völlig willkürlich, einer Zeit  $t'' > t'$  einen Wert  $\bar{\eta}'' < \bar{\eta}'$  zuzuschreiben, ebenso gut könnte man umgekehrt verfahren; denn welcher der beiden Zustände, der frühere oder der spätere, dem Grenzzustande näher liegt, ist allgemein nicht zu entscheiden. Es gelten eben hier dieselben Einwände wie gegen das *Boltzmannsche H-Theorem*, und es sei mir daher gestattet, auf die prinzipielle Seite dieser früher von mir erörterten Frage hier noch mit einigen Worten zurückzukommen.

Daß aus den Prinzipien der statistischen Mechanik, die auch ich für die einzig zuverlässige Grundlage solcher Wahrscheinlichkeitsbetrachtungen über dynamische Systeme halte, die beständige Zunahme irgend einer Phasenfunktion im Sinne des zweiten Hauptsatzes weder bewiesen noch auch wahrscheinlich gemacht werden kann, habe ich schon im Jahre 1896 mit Hilfe des *Liouvilleschen* und des *Poincaréschen* Satzes nachgewiesen<sup>1</sup> und diesen Beweis dann 1899 nach verschiedenen Richtungen ergänzt.<sup>2</sup> Es sei nämlich  $G$  irgend ein „invariantes Gebiet“ im Phasenraume, d. h. ein solches, aus welchem vermöge der Grundgleichungen weder Systeme austreten noch eintreten können,

<sup>1</sup> *Zermelo*, Ann. d. Phys. 57 (1896) S. 485.

<sup>2</sup> *Zermelo*, Phys. Ztschr. 1 (1900) S. 317.

general, is continuous. From the basic equation (21) of statistical mechanics cited above it follows with mathematical rigor that  $\eta = \log \frac{D}{N}$ , too, satisfies the differential equation

$$\frac{\partial \eta}{\partial t} = \sum \left( \frac{\partial \eta}{\partial p_i} \frac{\partial \varepsilon}{\partial q_i} - \frac{\partial \eta}{\partial q_i} \frac{\partial \varepsilon}{\partial p_i} \right)$$

and hence that  $\int \dots \int \eta dp_1 \dots dq_n$  is a motion invariant, i. e.,  $\bar{\eta} = \text{const.}$ , and this holds true entirely independently of the coordinate system, no matter whether we originally choose the volume elements greater or smaller, and determine the density-in-phase by means of one or the other passage to the limit.

But if a suitable reinterpretation of the concept of density had really led to the desired result, then it would certainly run counter to the basic equations given initially, i. e., to the entire theory previously developed. One would almost have to start from scratch and develop a new theory in which the theorem on the establishment of the statistical equilibrium could hold. But it is completely inadmissible to combine formulas of the old theory with the new theory, as the author does in the next chapter by adding, e. g., in (462) the assumption  $\bar{\eta}''_{12} \leq \bar{\eta}'_{12}$  to the consequences from previous developments. Of course, we can prove anything from contradictory assumptions, really proving only the inadmissibility of such a procedure. But I would like to raise a further objection against formula (462). Even if we assume that it is proved that  $\bar{\eta}$  converges with time to its minimal value, then this would have to hold true for decreasing just as for increasing times, as is conceded by the author himself (*Gibbs 1902*, p. 150, *Gibbs 1905*, p. 153), and the boundary values for  $t = +\infty$  and for  $t = -\infty$  would be the same. But then, it is entirely arbitrary to assign to a time  $t'' > t'$  a value  $\bar{\eta}'' < \bar{\eta}'$ , it is just as well to proceed conversely. For one cannot decide in general which of the two states, the earlier or the later one, is closer to the limit state. Here, the same objections apply as in the case of *Boltzmann's H-theorem*, and I may still be permitted a few remarks on the basic aspects of this question, which I discussed on an earlier occasion.

Already in 1896, I showed, using *Liouville's* and *Poincaré's* theorems,<sup>3</sup> that it is impossible to prove, or even make probable, the continual increase of any phase function in the sense of the second law from the principles of statistical mechanics, which I too consider the only reliable foundation of such probabilistic considerations about dynamical systems. In 1899 I provided various supplements to this proof.<sup>4</sup> For let  $G$  be any "invariant region" in the phase space, i. e., one that, because of the basic equations, systems can neither

<sup>3</sup> *Zermelo 1896a.*

<sup>4</sup> *Zermelo 1900.*

z. B. eines, das durch irgend welche Energiegrenzen  $\varepsilon'$  und  $\varepsilon''$  bestimmt ist, und es sei  $u$  irgend eine eindeutige Phasenfunktion. Dann wird der für das Gebiet  $G$  genommene Durchschnittswert in der *Gibbsschen* Bezeichnung

$$\bar{u} = \int u P dp_1 \dots dq_n$$

und ist von der Zeit unabhängig. Somit erhalten wir, weil die Phasenwahrscheinlichkeit  $P dp_1 \dots dq_n$  eine Bewegungsinvariante ist,

$$\frac{d\bar{u}}{dt} = \int \frac{du}{dt} P dp_1 \dots dq_n = \frac{d\bar{u}}{dt} = 0,$$

- 241 | d. h. die Funktion  $u$  wird im betrachteten Gebiete im Durchschnitt weder zunehmen noch abnehmen, sondern beides wird gleich wahrscheinlich sein. Dasselbe ergibt sich aber auch, wenn  $G$  kein invariantes Gebiet, sondern ein in zwei Grenzen  $u'$  und  $u''$  der betrachteten Phasenfunktion  $u$  eingeschlossener Bereich ist.<sup>1</sup> Daß ich in meiner damaligen Darstellung der Einfachheit halber  $P = 1$  angenommen hatte, ist für die Beweisführung selbst völlig unwesentlich. Meine Argumentation richtete sich nun auch gegen das *Boltzmannsche H-Theorem*, d. h. gegen die Behauptung, daß die von *Boltzmann* in die Gastheorie eingeführte Funktion  $H$  (*Boltzmann*, Vorlesungen über Gastheorie I § 5, 6), welche ebenso wie das *Gibbssche*  $\bar{\eta}$  im wesentlichen den Logarithmus der Phasenwahrscheinlichkeit darstellt und der Entropie entsprechen soll, beständig abnehmen müsse. In der sich hieraus entwickelnden Polemik<sup>2</sup> versuchte dann *Boltzmann* das *H-Theorem* zu retten, indem er seiner *H-Kurve* nicht mehr einen wesentlich absteigenden, sondern einen im allgemeinen gleichförmigen Verlauf parallel zur  $t$ -Achse zuschrieb, wobei die hin und wieder auftretenden „Buckel“ um so seltener sein sollen, je größer sie sind. Mag hiermit auch dem Minimalwerte von  $H$  entsprechend die überwiegende Wahrscheinlichkeit der *Maxwellschen* Verteilung plausibel gemacht werden, so kann ich doch in einem solchen nahezu symmetrischen Verlaufe der Kurve eine wirkliche Analogie zur Irreversibilität der Naturvorgänge keineswegs erblicken, und welchen Anspruch auf Allgemeingültigkeit dann die Ungleichung  $\frac{dH}{dt} \leq 0$  noch haben soll, ist mir unverständlich geblieben. Weder die *Boltzmannschen* noch die *Gibbsschen* Deduktionen haben meine Überzeugung erschüttern können, daß eine kinetische Wärmetheorie auch im Sinne der Wahrscheinlichkeitsrechnung sich nur dann mit dem zweiten Hauptsatze vereinigen lassen, wenn man sich entschließt, an Stelle der *Hamiltonschen* Bewegungsgleichungen solche Differentialgleichungen zugrunde zu legen, welche das Prinzip der Irreversibilität bereits in sich enthalten. Die Ausführungen des 13. Kapitels, in dem die Wirkung verschiedener äußerer Einflüsse auf Systeme und Gesamtheiten untersucht wird, beruhen

<sup>1</sup> *Zermelo*, Phys. Ztschr. 1 S. 319.

<sup>2</sup> *Boltzmann*, Ann. d. Phys. 57 (1896) S. 773; *ibid.* 60 (1897) S. 392. *Zermelo*, Ann. d. Phys. 59 (1896) S. 793, vgl. auch *Boltzmann*, Enzykl. d. math. Wiss. V. 8.

exit nor enter, such as one that is determined by any energy limits  $\varepsilon'$  and  $\varepsilon''$ , and let  $u$  be any single-valued phase function. Then, in *Gibbs's* terminology, the mean value taken for the region  $G$  becomes

$$\bar{u} = \int u P dp_1 \dots dq_n$$

and is independent of time. Thus, since the phase probability  $P dp_1 \dots dq_n$  is a motion invariant, we obtain

$$\frac{d\bar{u}}{dt} = \int \frac{du}{dt} P dp_1 \dots dq_n = \frac{d\bar{u}}{dt} = 0,$$

i. e., the function  $u$  will neither increase nor decrease on average in the region under consideration, but both cases will be equally probable. But we obtain the same result when  $G$  is not an invariant region but a domain enclosed within two boundaries  $u'$  and  $u''$  of the phase function  $u$  under consideration.<sup>5</sup> That I simply assumed that  $P = 1$  in my previous account for simplicity's sake is completely irrelevant to the proof itself. My line of argument was now also directed against *Boltzmann's*  $H$ -theorem, i. e., against the assertion that the function  $H$  introduced into the theory of gases by *Boltzmann* (1896a, § 5, 6), which, like *Gibbs's*  $\bar{\eta}$ , basically represents the logarithm of the phase probability and is supposed to correspond to the entropy, would have to decrease continually. In the ensuing polemic<sup>6</sup> *Boltzmann* then tried to salvage the  $H$ -theorem by attributing to his  $H$ -curve a course that is no longer essentially descending but generally uniform and parallel to the  $t$ -axis, where the occasionally occurring "bumps" are supposed to be the rarer the greater they are. While this may well make plausible the predominant probability of the *Maxwell* distribution in accordance with the minimal value of  $H$ , I certainly cannot recognize a genuine analogy to the irreversibility of the natural processes in such an almost symmetric course of the curve, and I fail to see how then the general validity of the inequality  $\frac{dH}{dt} \leq 0$  is supposed to be justified. Neither *Boltzmann's* nor *Gibbs's* deductions have been able to shake my conviction that it is possible to unite a kinetic theory of heat also in the sense of probability calculus with the second law only if we resolve to replace as a basis *Hamilton's* equations of motion by differential equations already containing the principle of irreversibility in themselves.

The elaborations of chapter 13, which examines the effects of various external influences on systems and ensembles, rest for the most part on the

<sup>5</sup> *Zermelo 1900*, p. 319.

<sup>6</sup> *Boltzmann 1896*; *Boltzmann 1897*. *Zermelo 1896b*, see also *Boltzmann and Nabl 1907*.

nun größtenteils auf den im 12. Kapitel gewonnenen Ergebnissen und sind daher nach den vorstehenden Ausführungen nicht frei von inneren Widersprüchen. Es wird angenommen, daß eine Gesamtheit, wenn einmal durch äußere Einflüsse ihr statistisches Gleichgewicht gestört ist, innerhalb einer gewissen Zeit wieder zu einem solchen Gleichgewichte zurückkehren werde, und hieraus werden dann in Verbindung mit den sonst abgeleiteten Beziehungen Ungleichungen gewonnen, die den thermodynamischen zu entsprechen scheinen. Im letzten Kapitel werden endlich die allgemeinen Betrachtungen spezialisiert auf Systeme von Molekülen verschiedener Gattungen, in welchen gegenseitige Umsetzungen zwischen den verschiedenen Gattungen stattfinden können. Auch solche Gesamtheiten denkt *Gibbs* sich kanonisch verteilt mit einem Wahrscheinlichkeitsexponenten der Form

$$H = \frac{\Omega + \mu_1\nu_1 + \dots + \mu_h\nu_h - \varepsilon}{\theta}, \quad (503)$$

242 | wo die Zahlen  $\nu_1, \nu_2, \dots, \nu_h$  Anzahlen für die verschiedenen Partikelgattungen und  $\Omega, \theta, \mu_1, \mu_2, \dots, \mu_h$  Konstanten bezeichnen. Aus diesem Ansatz ergeben sich dann wieder Gleichgewichtsbedingungen wie die folgende

$$d\Psi = \bar{H} d\theta + \sum \mu_1 d\nu_1 - \sum \bar{A}_1 da_1, \quad (532)$$

die der thermodynamischen analog ist, wenn  $\theta$  der Temperatur,  $-\bar{H}$  der Entropie und  $\Psi = \bar{\varepsilon} + \theta\bar{H}$  der freien Energie entspricht. Es ist aber physikalisch nicht recht einzusehen, wie denn das chemische Gleichgewicht der einzelnen Systeme mit dem statistischen Gleichgewichte ganzer Gesamtheiten zusammenhängen soll. Auch beruht die formale Analogie mit den thermodynamischen Gleichungen zum großen Teile nur auf der zu diesem Zwecke gewählten Form des Wahrscheinlichkeitsexponenten. Ich kann daher auch hier nicht zugeben, daß wir von solchen Analogien einen tieferen Einblick in das Wesen der chemischen Umsetzungen und ihrer Gleichgewichtsbedingungen zu erwarten hätten.

Einige weitere in dem besprochenen Werke mehr nebensächlich behandelte Gegenstände will ich zum Schlusse noch kurz erwähnen: im 2. Kapitel die Anwendung des Gesetzes von der Erhaltung der Phasenausdehnung auf die „Fehlertheorie“, wo aber im wesentlichen nur der Begriff, nicht die „Erhaltung“ der Phasenausdehnung benutzt wird, die Frage also nicht eigentlich der statistischen Mechanik angehört, und im 3. Kapitel die Anwendung desselben Prinzips auf die Integration der Bewegungsgleichungen mit Hilfe des „letzten Multiplikators“, für die aber statt *Jacobis* „Vorlesungen über Dynamik“ eine viel spätere Arbeit von *Boltzmann* als Quelle zitiert wird. Besonders verdienstvoll sind die Untersuchungen des 7. Kapitels, wo die Berechnung einer großen Zahl von Durchschnittswerten ausgeführt und Methoden zu ihrer Fortsetzung angegeben werden. Hervorzuheben ist dabei die eingehende Berücksichtigung der „mittleren Fehler“, deren Abschätzung allein uns gelegentlich berechtigen kann, die in der Untersuchung vorkommenden Größen

results obtained in chapter 12, and hence are, according to the previous remarks, not devoid of inner contradictions. It is assumed that an ensemble, when once disturbed in its statistical equilibrium by external influences, returns to such an equilibrium within a certain period of time, and from this, in connection with the relations otherwise derived, inequalities are obtained that appear to correspond to the thermodynamical ones. Finally, in the last chapter, the general considerations are specifically applied to systems of molecules of various categories, in which mutual conversions among the various categories take place. *Gibbs* assumes that such ensembles, too, are canonically distributed with an index of probability of the form

$$H = \frac{\Omega + \mu_1\nu_1 + \dots + \mu_h\nu_h - \varepsilon}{\theta}, \quad (503)$$

where the numbers  $\nu_1, \nu_2, \dots, \nu_h$  denote numbers for the various categories of particles, and  $\Omega, \theta, \mu_1, \mu_2, \dots, \mu_h$  denote constants. From this ansatz in turn follow equilibrium conditions, such as

$$d\Psi = \bar{H} d\theta + \sum \mu_1 d\bar{\nu}_1 - \sum \bar{A}_1 da_1, \quad (532)$$

which is analogous to the thermodynamical one, if  $\theta$  corresponds to the temperature,  $-\bar{H}$  to the entropy, and  $\Psi = \bar{\varepsilon} + \theta\bar{H}$  to the free energy. But from a physical perspective it is difficult to see how the chemical equilibrium of the individual systems should be related to the statistical equilibrium of entire ensembles. Moreover, the formal analogy with the thermodynamical equations rests for the most part only on the form of the index of probability chosen for this purpose. Thus here, too, I cannot concede that we should expect to gain deeper insights from such analogies into the nature of the chemical conversions and their equilibrium conditions.

In closing, I would like to mention some further matters which did not receive full attention in the work under review: the application of the law of the preservation of the extension-in-phase to the “error theory” in chapter 2, where, however, essentially only the concept, and not the “preservation”, of the extension-in-phase is used and thus where the question does not really lie within the purview of statistical mechanics, and the application of the same principle to the integration of the equations of motion by means of the “last multiplier” in chapter 3, for which, however, instead of *Jacobi’s* “Vorlesungen über Dynamik”, a much later work by *Boltzmann* is cited as a source. Particularly commendable are the investigations in chapter 7, where a great number of mean values are calculated and methods for their continuation are presented. I should highlight here the close attention paid to the “mean errors” whose estimation alone can, on occasion, justify the statistical replacement

statistisch durch ihre Durchschnittswerte zu ersetzen. Interessant ist auch das 11. Kapitel, in dem eine Reihe von Sätzen über Maximums- und Minimumseigenschaften verschiedener Verteilungen entwickelt werden. —

In seinem Vorworte bekennt der Verfasser, er versuche nicht, die Geheimnisse der Natur zu erklären, sondern wolle sich mit dem bescheideneren Ziele begnügen, die einfachsten Sätze einer Disziplin zu entwickeln, die einer künftigen mechanischen Naturerklärung als Grundlage dienen könne. Auf diesem Gebiete sei auch schon jetzt ein sicheres Fortschreiten möglich, während einer eingehenden Durchführung der Molekularhypothese zur Zeit noch unüberwindliche Schwierigkeiten entgegenständen. Hat ihn bei der Verfolgung dieses Zieles sein lebhaftes Interesse für die Sache auch gelegentlich verleitet, die selbstgesteckten Grenzen zu überschreiten und dem Irrtum zu verfallen, so kann doch das Unternehmen selbst, mit dem der hochverdiente Forscher seine Lebensarbeit beschlossen hat, nur als gelungen betrachtet werden. Wo immer man statistischer und wahrscheinlichkeitstheoretischer Betrachtungen in der Mechanik bedarf, wird man nicht umhin können, sich zunächst mit dem *Gibbs*schen Werke vertraut zu machen und auf den hier gelegten Grundlagen weiterzubauen.

Göttingen.

*E. Zermelo.*

of the magnitudes occurring in the investigation by their mean values. Of further interest is chapter 11, where a series of theorems on maximum and minimum properties of various distributions is developed.—

In his preface, the author declares that he will not seek to explain the secrets of nature but content himself with the more modest goal of developing the simplest theorems of a discipline that may serve as the basis for a future mechanistic explanation of nature. [He states that] while, already today, it is possible to make sound progress in this field, a detailed implementation of the molecular hypothesis still faces unsurmountable difficulties. Even if in the pursuit of this goal his lively interest in the matter has, on occasions, led him to cross the self-imposed boundaries and to fall into error, the enterprise itself, which marks the conclusion of the life-work of this exemplary researcher, can only be considered a success. Anyone in need of statistical and probabilistic considerations in mechanics cannot fail to first turn to *Gibbs's* work and to build on the foundations laid in it.

Göttingen.

*E. Zermelo.*