

Essay I: A Critical Examination of the Statistical Concept of Entropy in Classical Physics

Summary. The first sections (§§2–3) contain preliminary discussions. For a simple classification system, the following concepts are defined: degree of order o^* and degree of confirmation c^* (§2); degree of disorder d^* and entropy S^* (§3). The main body of Essay I deals with the concept of entropy in classical statistical mechanics and its relation to the thermodynamic concept of entropy (S_{th}). First, Boltzmann's definition of entropy (S_B , §4) and his H -theorem (§5) are explained. S_B applies only to statistical descriptions (D^{st}). Two methods for extending it to other forms of descriptions, especially individual descriptions (D^{ind}), are discussed (§6). It is shown that the customary extended concept S_B^{II} is not generally in agreement with S_{th} and is a logical rather than a physical concept. An alternative extended concept S_B^{I} is proposed, which is a physical concept and is generally in agreement with S_{th} . Gibbs's method of representing a description of a state of a gas by an ensemble in the γ -space is explained (§7). Gibbs's entropy concept S_G^{II} is examined (§8); it is similar to S_B^{II} and has the same disadvantages. It is shown that Gibbs's theorem to the effect that S_G^{II} increases in the course of time is not an analogue to the second law of thermodynamics. An alternative S_G^{I} for Gibbs's S_G^{II} is defined which avoids the disadvantages of the latter (§9). A theorem on S_G^{I} is formulated which is an analogue of the second law. Some authors equate entropy with negative amount of information. For D^{st} , the two values are indeed equal. However, the relation holds for other descriptions only if S_B^{II} , not S_B^{I} , is chosen, hence only if the physical character and the agreement with S_{th} are sacrificed (§10).

A Brief Formulation of My Main Point Concerning the Statistical Concept of Entropy

A. GENERAL REMARKS ON MACRO- AND MICRO-MAGNITUDES.

Let W be a physical macro-magnitude (e.g., volume or mass of a body), for which a measuring procedure is established which, according to experience, yields in any given case (of the kind to which W is applicable) a definite value of W independently of subjective factors of the observer (e.g., his preferences or prior information). Then we may say that W is an *objective* concept. (I shall later, in §6, use the more specific term ‘purely physical magnitude’ and give more exact criteria for it.) For the question of the objectivity of a magnitude W it makes no difference whether the value of W is directly measured or indirectly determined, i.e., calculated according to established rules from directly measured values of other magnitudes. If the latter magnitudes are objective physical concepts, then W is likewise.

According to a basic assumption of classical physics (usually not formulated explicitly) the result of a measurement of W for a physical system s at a given time t_1 depends upon the micro-state of s at t_1 . Let us formulate this assumption more exactly. The process of measurement does actually extend over a certain period of time around t_1 , say Δ_1 , and covers a certain spatial region R containing s . The assumption means that, for any objective physical macro-magnitude W there is a *corresponding micro-magnitude M* , i.e., one which fulfills the following conditions:

- (1) There are some of the *fundamental micro-magnitudes* (e.g., electro-magnetic field, gravitational field, electric charge, mass, position of particles, etc.) which are relevant for W , say M_1, \dots, M_r, \dots .
- (2) The value of M for any region R and any period Δ can be defined as a function of the distribution of the magnitudes M_i within R during Δ .

(3) If W is applicable to the state of s during Δ_1 , then $W(s, \Delta_1) = M(R, \Delta_1)$. (This agreement between M and W is either meant to hold exactly or to be such that the value of W lies with high probability within a small interval around the value of M .)

Since the fundamental micro-magnitudes of classical physics were conceived as objective, M is likewise objective. (The relativity of a physical magnitude with respect to a basis of reference, e.g., zero point, unit, scale form, coordinate system or the like must not be misunderstood as subjectivity.) This means that the value of M is

dependent only on the objective micro-situation, not on the information possessed by the observer X . (Needless to say, X 's knowledge of M , that is, his estimate of the actual value of M , depends upon his information concerning observational results.)

Let D_0 be the complete description of the distribution of the values of the magnitudes M_i within R during Δ_1 . Let $D_0, D_1, \dots, D_n, \dots$ be a series of descriptions of R during Δ_1 , such that D_{n+1} is weaker than D_n (i.e., D_{n+1} is deducible from D_n but not vice versa) and such that each description D_n of the series is sufficient for the determination of M for R and Δ_1 . Let the value of M derived from D_n be $M(D_n)$. Obviously the values $M(D_0), M(D_1), \dots$ must all be equal, because all these descriptions, being deducible from the true complete description D_0 , are likewise true and hence cannot lead to a value of M different from the actual value of M for R and Δ_1 as given by $M(D_0)$.

B. APPLICATION TO ENTROPY.

We shall now apply the above considerations to thermodynamic entropy (S_{th}), a macro-magnitude of classical thermodynamics. For simplicity, we consider the state of an isolated gas body g during the time period Δ_1 around t_1 . Let the vessel containing the gas be divided into two chambers separated by a wall. To assure the applicability of S_{th} , we assume that the gas is during Δ_1 in a state of local thermodynamic equilibrium (§5); this means that the temperature is the same throughout each chamber and throughout Δ_1 , and the same holds for the pressure. Temperature and pressure are directly measurable macro-magnitudes. The entropy S_{th} is not immediately measurable. Its value for g at t_1 (relative to some final state of g , say, the state of total equilibrium with temperature and pressure equalized throughout the vessel) can be determined by a series of measurements during a slow (quasi-reversible) transition process leading from the given state to the chosen final state. On the other hand, S_{th} for g at t_1 can also be calculated with the help of the laws of thermodynamics from the following data: the kind of gas (say, hydrogen), and for each chamber its volume and the values of temperature and pressure of the gas in the chamber at t_1 . Thus S_{th} can be determined on the basis of measurable, objective macro-magnitudes; hence S_{th} is itself an objective macro-magnitude.

The micro-theory for thermal processes in gases is the kinetic theory of gases which applies the methods of statistical mechanics to the molecules of the gas. Thus the problem arises of defining in

terms of this theory statistical micro-magnitudes corresponding to the thermodynamic macro-magnitudes, in particular, to S_{th} . It is found that the fundamental micro-magnitudes relevant for the thermodynamic macro-magnitudes are the following: the mass of a molecule (which is the same for all molecules of a given substance), and for each molecule at a given time certain state-coordinates (in the simplest case, that of a sufficiently diluted monatomic gas, there are six of these, viz., the three position coordinates and the three momentum components). Thus the task is to define on this basis a suitable statistical concept of entropy.

Solutions of this problem were proposed by Boltzmann and Gibbs. Boltzmann applied his concept S_B only to descriptions of a certain kind (I shall call them statistical descriptions, D^{s}) which are much weaker than complete micro-descriptions D_0 . Later physicists proposed a generalization of S_B , which I shall denote by ' S_B^{II} '; this concept is applicable to descriptions of various series D_0, D_1, D_2 , etc., as described above. We find that the resulting values of S_B are not exactly equal for any two of the descriptions, and show for some of them even enormous differences, one value being a large multiple of the other. Since all these descriptions are true micro-descriptions of the same gas g for the same period Δ_1 around t_1 , at most one of the resulting values of S_B^{II} can exactly agree with the value of S_{th} for g at t_1 as determined from measurements made during Δ_1 , and some will diverge widely from S_{th} .

The statistical entropy concept proposed by Gibbs, which I shall denote by ' S_G^{II} ', has the advantage of being applicable to a much wider class of cases than Boltzmann's concept and to a greater variety of logical forms of descriptions. But S_G^{II} shares with Boltzmann's concept the feature just explained; its value varies in the given series of descriptions. Therefore S_G^{II} cannot possibly agree generally with S_{th} . This is my main objection against both S_B^{II} and S_G^{II} . It seems the generally accepted view among physicists that S_B^{II} and S_G^{II} are adequate micro-magnitudes corresponding to S_{th} . I believe that the facts stated refute this view.

The variation of both concepts in the above-mentioned series of descriptions shows also that both concepts are not objective. Today many physicists seem to recognize this. I shall show, moreover, that either concept is purely logical; that is to say, a statement which ascribes to g during Δ_1 a specific value of either S_B^{II} or S_G^{II} does not give any information at all concerning the physical state of g . This is my second main objection to the view that these concepts correspond to S_{th} . The result just stated follows from the fact that the value of either magnitude calculated on the basis of a micro-

description, depends merely upon the strength or *amount* of information of the description, not on its *content*; in other words, the value is determined merely by *how much* the description says, not by *what* it says.

Some physicists today believe that the non-objective (or not purely objective) character is inevitable for any statistical concept of entropy. (Incidentally, some believe the same even for a statistical concept of temperature.) This view is in conflict with the principle, implicitly accepted by physicists in the classical period, that for any objective macro-magnitude there must be a corresponding objective micro-magnitude. In defense of this principle I shall propose an alternative S_B^I to Boltzmann's S_B^{II} and an alternative S_G^I to Gibbs's S_G^{II} such that the following conditions are fulfilled if S is either S_B^I or S_G^I :

(1) S is an objective physical micro-magnitude; that is, a statement of a specific value of S for g during Δ_1 gives some information about the physical state of g .

(2) The value of S is the same for all descriptions of the series D_0, D_1, \dots .

(3) The value of S determined on the basis of any of these descriptions is in agreement with the value of S_{th} for g during Δ_1 . (The definition of S_B^I is constructed in a very simple way. Since it is known that for any D^{st} , S_B^{II} is in agreement with S_{th} , I define S_B^I in such a way that it has for any D^{st} the same value as S_B^{II} , and then I assign to it for any other description of the same series this same value. Thus the three conditions are fulfilled. The guiding idea for the definition of S_G^I is analogous but somewhat more complicated.)

2

A System of Qualitative Classification

Summary. In a classification of K cells Q_j ($j = 1, \dots, K$), an individual description (D^{ind}) assigns each of the N elements to one cell. A statistical description (D^{st}) states merely the number N_j of elements of each cell Q_j . Z is the number of all D^{ind} ; z is the number of those D^{ind} which correspond to a given D^{st} . The degree of order o^* of a D^{ind} is defined as $1/z$. A measure function m^* for the D^{ind} is defined, which is proportional to o^* . With its help, the degree of confirmation c^* is defined.

As a preliminary to the study of systems with quantitative magnitudes, we shall consider in this and the next section the much simpler case of classification systems. Here the N elements are characterized by qualitative properties Q_j ($j = 1, \dots, K$) which form a classification or division (i.e., each element has one and only one of the K properties). The K properties Q_j are also called the *cells* of the classification. A complete description of the system specifies for each element the cell to which it belongs. A description of this kind is called an *individual description* D^{ind} .

(2-1) The number of individual descriptions for given N and K is $Z = K^N$.

Throughout this section, ' D ' without superscript (with or without subscript) refers to a D^{ind} . For a given D_i , let N_j ($j = 1, 2, \dots, K$) be the number of elements belonging to the cell Q_j . The numbers N_1, \dots, N_K are called the *cell numbers* in D_i . D_1 and D_2 are said to be *statistically equivalent* if they have the same cell numbers.

A description of the system of N elements which states for each cell Q_j merely its cell number N_j (but does not say which elements belong to Q_j) is called a *statistical description* D^{st} .

(2-2) (a) The number of D^{ind} with the cell numbers N_1, \dots, N_K is

$$z = \frac{N!}{N_1! \dots N_K!}.$$

In other words:

(2-2) (b) Let D_i^{ind} be an individual description with the cell numbers N_1, \dots, N_K and D_i^{st} be the corresponding statistical description. Let $z(D_i^{\text{ind}})$ be the number of D^{ind} statistically equivalent to

D_i^{ind} , and $z(D_i^{\text{st}})$ the number of D^{ind} corresponding to D_i^{st} ; then

$$z(D_i^{\text{ind}}) = z(D_i^{\text{st}}) = \frac{N!}{N_1! \dots N_K!}.$$

(2-3) z has its *minimum* in the case that all elements belong to the same cell ("homogeneous system"), hence one cell number is N and all others 0:

$$z_{\min} = 1.$$

(2-4) z has its *maximum* in the case that all cell numbers are as nearly equal as possible (i.e., the difference between any two is 0 or 1). Let us assume, for simplicity, that N is divisible by K ; then

$$z_{\max} = \frac{N!}{\left(\frac{N}{K}\right)!^K}.$$

The traditional concept of regularity or uniformity is often used in philosophical discussions, e.g., in the postulate of the uniformity of the world as an allegedly necessary presupposition of induction. This concept is meant in such a sense that one system (possible world) is regarded as more regular than another if more or stronger laws (either universal or statistical) hold in the one than in the other. Let us try to explicate this concept, at least for certain cases, in a quantitative form as "degree of order" of a D^{ind} , denoted by ' $o(D^{\text{ind}})$ '. At another place (not yet published) I have analyzed in greater detail the traditional concept mentioned and the conditions which an adequate explicatum o must fulfill. I have shown that with respect to simple classification systems, the function o must fulfill the following requirements:

(2-5) *First requirement.* For given N and K , $o(D)$ depends only on the cell numbers; in other words, if D_1 and D_2 are statistically equivalent, $o(D_1) = o(D_2)$.

(2-6) *Second requirement.* $o(D)$ is a symmetric function of the cell numbers in D .

(2-7) *Third requirement.* Let D_2 be formed from D_1 by shifting an element from one cell to another cell which had in D_1 the same or a higher cell number than the first cell. Then

$$o(D_2) > o(D_1).$$

It is easily seen that, among others, the following function o^* fulfills the requirements mentioned:

$$(2-8) \quad o^*(D) =_{\text{Def}} \frac{N_1! \dots N_K!}{N!} = \frac{1}{z(D)}.$$

Once a suitable function o is found, it seems plausible (as I have explained at the other place) to define on its basis the *degree of confirmation* in the following way. The degree of confirmation $c(h, e)$ of a hypothesis h with respect to the evidence e is definable in terms of a measure function m (expressing the initial inductive probability or the degree of confirmation on the null evidence) as follows ((Prob.) §55A):

$$(2-9) \quad c(h, o) =_{\text{Def}} \frac{m(e, h)}{m(e)}.$$

Now it seems plausible to choose the function m in such a way that $m(D)$ is proportional to $o(D)$. The factor of proportionality is uniquely determined by the condition that the sum of m for all D^{ind} must be 1. Then we take, in the usual way, $m(h)$ for any statement h as the sum of m for those D^{ind} in which h holds ((Prob.) §55A). Applying this procedure to the function o^* defined by (2-8), we obtain the following function m^* :

$$(2-10) \quad m^*(D) =_{\text{Def}} \frac{N_1! \dots N_K!(K-1)!}{(N+K-1)!}.$$

Then we define:

$$(2-11) \quad c^*(h, e) =_{\text{Def}} \frac{m^*(e, h)}{m^*(e)}.$$

The same functions m^* and c^* , derived in a different way without the use of degree of order, were suggested by me (first in 1945) as a basis of a system of inductive logic ((Prob.) §10).

Let e be a D^{ind} for the elements a_1, \dots, a_n , and n_j the cell number of Q_j in e . Let h_j be the hypothesis that another element a_{n+1} belongs to Q_j . The following result is easily derived from (2-10) and (2-11):

$$(2-12) \quad c^*(h_j, e) = \frac{n_j + 1}{n + K}.$$

3

Disorder and Entropy for a Classification System

Summary. The degree of disorder d^* is defined as $1/o^* = z$. In analogy to Boltzmann's definition, we define the entropy S for a classification system as a linear function of the logarithm of $d^*(=z)$. We choose a particular function S^* of this form, defined as $\text{Log}(a/z)$ ('Log' denotes the logarithm to base 2).

For the further discussions in Part I, we shall leave aside the problems of inductive logic, to which we shall return in Part II. Our main topic in Part I will be the problem of entropy. The aim of the present section is to define disorder and entropy for classification systems. If any function o is chosen as degree of order, let us take as *degree of disorder* d its reciprocal:

$$(3-1) \quad d(D) = \text{df} \frac{1}{o(D)}.$$

We apply this to o^* (2-8) and define:

$$(3-2) \quad \begin{aligned} \text{(a)} \quad d^*(D) &= \text{df} \frac{N!}{N_1! \dots N_K!} \\ \text{(b)} \quad &= z(D). \end{aligned}$$

Since a given D_i^{st} is characterized by the set of cell numbers which holds for each of the corresponding D^{ind} , and since d^* depends merely on these cell numbers, it seems natural to ascribe that value of d^* which those D^{ind} have in common to D_i^{st} also. Therefore we define:

$$(3-3) \quad \begin{aligned} \text{(a)} \quad d^*(D_i^{\text{st}}) &= \text{df} \frac{N!}{N_1! \dots N_K!} \\ \text{(b)} \quad &= z(D_i^{\text{st}}). \end{aligned}$$

Hence:

(3-4) If D_i^{st} corresponds to D_i^{ind} ,

$$d^*(D_i^{\text{st}}) = d^*(D_i^{\text{ind}}).$$

From (2-3):

$$(3-5) \quad d_{\min}^* = 1.$$

From (2-4):

$$(3-6) \quad d_{\max}^* = \frac{N!}{\left(\frac{N}{K}\right)!^K}.$$

In order to obtain for a classification system a concept S of *entropy* analogous to Boltzmann's concept (§4), we shall define S as a linear function of the logarithm of the degree of disorder d^* , which is equal to z . (We write 'ln' for the natural logarithm.) Since Boltzmann defined his concept only for D_i^{st} , we shall at present do the same. Later, however, we shall apply our concept also to D^{ind} .

Thus for given N and K and for a given D_i^{st} , $S(D_i^{\text{st}})$ is obtained by multiplying $\ln z(D_i^{\text{st}})$ by a constant coefficient (>0) and adding another constant. The two constants may depend upon N and K . Therefore we shall represent them as functions $f(N, K)$ and $g(N, K)$, respectively:

(3-7) $S(D_i^{\text{st}})$ shall have the form $f(N, K) \ln z(D_i^{\text{st}}) + g(N, K)$ with suitably chosen functions f and g .

We wish the entropy to be a monotone increasing function of disorder; therefore

(3-8) For any N and K , $f(N, K) > 0$.

With the help of Stirling's theorem, we have the following approximation, where $r_i = N_i/N$:

$$(3-9) \quad (a) \quad \ln z = \ln \frac{N!}{\prod_j N_j!} \cong N \sum_j \left(r_j \ln \frac{1}{r_j} \right) + \frac{1}{2} \left(\ln N - \sum_j \ln N_j \right) - (K-1) \ln \sqrt{2\pi},$$

$$(b) \quad \cong N \sum_j \left(r_j \ln \frac{1}{r_j} \right).$$

We lay down the convention that the value of $u \ln u$ (and hence also of $u \ln 1/u$) for $u=0$ is to be taken as 0 (since this is the limit of the function for $u \rightarrow 0$). The cruder approximation (3-9)(b) is applicable only if N is very large; its use is customary in statistical mechanics. From (2-3):

$$(3-10) \quad \ln z_{\min} = 0.$$

From (2-4) with (3-9):

$$(3-11) \text{ (a)} \quad \ln z_{\max} \cong \left(N + \frac{K}{2} \right) \ln K - \frac{K-1}{2} \ln N - (K-1) \ln \sqrt{2\pi},$$

$$\text{(b)} \quad \cong N \ln K, \text{ if } N \text{ is very large in relation to } K.$$

Applying (3-9)(b) to (3-7):

(3-12) For large N , $S(D_i^{\text{st}})$ shall have approximately the form

$$F(D_i^{\text{st}}) = F(N, K, N_1, \dots, N_K) = N \sum \left(r_j \ln \frac{1}{r_j} \right) f(N, K) + g(N, K).$$

We are free to choose the functions f and g in (3-7). We shall now lay down two requirements which seem desirable for our purposes, and then we shall choose as f and g two functions f_2 and g_2 , respectively, such that these requirements are fulfilled for the form F in (3-12). Both requirements can be fulfilled for the form (3-7) only as approximations for large N ; but the functions f_2 and g_2 thus determined will later be used in the definition of S generally, for any N .

(3-13) *First requirement.* If N varies but the relative frequencies r_j remain unchanged, F is to be proportional to N .

This means the following. Let D be a D^{st} characterized by N , K , and the values N_j . Let D' be analogous to D but with N' , K' , and N'_j such that $N' = mN$ with a fixed positive integer m , $K' = K$, and for every j , $N'_j = mN_j$, hence $r'_j = N'_j/N' = r_j$. Then, the requirement says that $F(D') = mF(D)$. Hence from (3-12) (with ' $\sum(\dots)$ ' for ' $\sum_j \left(r_j \ln \frac{1}{r_j} \right)$ '):

$$mN \sum (\dots) f(mN, K) + g(mN, K) = mN \sum (\dots) f(N, K) + mg(N, K).$$

$$mN \sum (\dots) [f(mN, K) - f(N, K)] = mg(N, K) - g(mN, K).$$

If, for given N , K and m , the values N_j vary, \sum varies but everything else remains constant. Therefore the factor of \sum must be 0, and hence also the right side. Thus, for any N , K , m , $f(mN, K) = f(N, K)$ and $g(mN, K) = mg(N, K)$. Therefore:

(3-14) $f(N, K)$ is independent of N ; we write ' $f_1(K)$ '.

The condition for g just mentioned is fulfilled if we put

$$(3-15) \quad g_1(N, K) = {}_{Df} N \cdot h(K),$$

with an arbitrary function h . Hence:

$$(3-16) \quad F(\dots) = N \sum (\dots) f_1(K) + Nh(K).$$

(3-17) *Second requirement.* If each cell is subdivided into a fixed number p of new cells with equal cell numbers N_j/p , F is to remain unchanged.

Here we consider two D^s , D and D' , with the following relations. The number of elements is the same in both: $N' = N$. Each cell in D corresponds to p cells in D' with equal cell numbers: $N_j = N'_{j1} + N'_{j2} + \dots + N'_{jp}$; $N'_{j1} = N'_{j2} = \dots = N'_{jp} = N_j/p$. Hence $r'_{j1} = \dots = r'_{jp} = r_j/p$. Corresponding to each term in \sum in (3-16) for D , there are for D' p equal terms $\frac{r_j}{p} \ln \frac{p}{r_j}$; their sum is $r_j \left(\ln \frac{1}{r_j} + \ln p \right)$.

Thus we have, since $\sum r_j = 1$:

$$(3-18) \quad F(D') = N \left[\sum \left(r_j \ln \frac{1}{r_j} \right) + \ln p \right] f_1(pK) + Nh(pK).$$

According to (3-17), the right sides in (3-16) and (3-18) must be equal; hence:

$$\sum (\dots) [f_1(pK) - f_1(K)] + \ln(p) \cdot f_1(pK) + h(pK) - h(K) = 0.$$

As in the earlier case, the factor of \sum must be 0; hence $f_1(pK) = f_1(K)$. Thus f_1 must be a positive constant (see (3-8)), which we denote by ' A ':

$$(3-19) \quad f_2(N, K) = {}_{Df} A (> 0).$$

Then

$$A \ln(p) + h(pK) - h(K) = 0.$$

We put $h'(K) = {}_{Df} e^{h(K)} > 0$. Hence $h(K) = \ln h'(K)$. Therefore

$$\ln h'(pK) = \ln h'(K) - A \ln p = \ln [h'(K)p^{-A}].$$

Hence $h'(pK) = h'(K)p^{-A}$. Let $h'(K)$ have the form $\sum_n (C_n K^{E_n})$,

where the exponents E_n and the corresponding coefficients C_n may be any real numbers. Then:

$$\sum (C_n p^{E_n} K^{E_n}) = \sum (C_n p^{-A} K^{E_n}).$$

Since this must hold for any K , coefficients of equal powers of K must be equal:

$$\text{For any } n, C_n p^{E_n} = C_n p^{-A}.$$

For any n , this condition is fulfilled if and only if either $C_n = 0$ or $p^{E_n} = p^{-A}$ and hence, for $p > 1$, $E_n = -A$. Therefore, for $E_n \neq -A$, all coefficients $C_n = 0$; for $E_n = -A$, C_n is arbitrary. Hence $h'(K) =_{\text{Df}} CK^{-A}$, with arbitrary C . Since $h'(K) > 0$, $C > 0$. We put $B = \ln C$. Then $h(K) =_{\text{Df}} B - A \ln K$; and from (3-15),

$$(3-20) \quad g_2(N, K) =_{\text{Df}} N(B - A \ln K).$$

Thus our result is as follows:

(3-21) If $F(D_i^{\text{st}})$ is defined in the form

$$N \left\{ A \left[\sum \left(r_j \ln \frac{1}{r_j} \right) - \ln K \right] + B \right\}$$

with arbitrary numerical constants $A (> 0)$ and B , then F fulfills the requirements (3-13) and (3-17).

A function S of the form (3-7) has for large N approximately the form F (3-12). Therefore we decide to use in (3-7) those functions f_2 and g_2 which have been found to fulfill the requirements:

(3-22) $S(D_i^{\text{st}})$ shall have the form

$$A[\ln z(D_i^{\text{st}}) - N \ln K] + BN,$$

with constants $A (> 0)$ and B to be chosen.

Any function S of this form fulfills the two requirements if N is large. The definition of S will not, of course, be restricted to large N ; but the requirements cannot be fulfilled for all N by any function of the form (3-7); this can be seen by applying the requirements to the form (3-9)(a) instead of (3-9)(b).

If we had only to do with classifications, as in this section, the second requirement would not appear as essential, and therefore a simpler definition of S without the term ' $-N \ln K$ ' would presumably appear as preferable. However, we introduce the concept of

entropy for a classification chiefly as a model to guide our later construction of the abstract entropy concept. This concept will involve quantitative magnitudes, and therefore the invariance with respect to a subdivision of the cell system will be important.

We shall now define the function S^* as a special case of the form (3-22). We choose $B = 0$. Further, we shall use the logarithm to the base 2 (denoted by 'Log') because we intend to use S^* , together with d^* , o^* , m^* , and c^* , in inductive logic and in the theory of information, where the base 2 is convenient and customary. This makes it convenient to choose $A = 1/\ln 2$. Thus we define:

$$(3-23) \quad S^*(D_i^{st}) = \text{Log}_2 z(D_i^{st}) - N \text{Log} K.$$

(3-24) S^* fulfills the two requirements (3-13) and (3-17) approximately for large N .

With (2-1):

$$(3-25) \quad S^*(D_i^{st}) = \text{Log} \frac{z(D_i^{st})}{Z}.$$

$z(D_i^{st})$ is the proportion of those D^{ind} which correspond to the given D_i^{st} ; hence it is $< Z$. Therefore:

(3-26) For any D^{st} , $S^* < 0$.

According to (3-3), the relation between S^* and d^* is as follows:

$$(3-27) \quad S^*(D_i^{st}) = \text{Log} d^*(D_i^{st}) - N \text{Log} K.$$

With (3-9):

$$(3-28) \quad \begin{aligned} (a) \quad S^*(D_i^{st}) &\cong N \left[\sum_i \left(r_j \text{Log} \frac{1}{r_j} \right) - \text{Log} K \right] \\ &\quad + \frac{1}{2} \left(\text{Log} N - \sum_j \text{Log} N_j \right) - (K-1) \text{Log} \sqrt{2\pi}, \\ (b) \quad &\cong N \left[\sum_i \left(r_j \text{Log} \frac{1}{r_j} \right) - \text{Log} K \right], \text{ for large } N. \end{aligned}$$

With (3-10) and (3-11):

$$(3-29) \quad S_{\min}^* = -N \log K.$$

$$(3-30) \quad (a) \quad S_{\max}^* \cong - \left[\frac{K-1}{2} \log N - \frac{K}{2} \log K + (K-1) \log \sqrt{2\pi} \right]$$

$$(b) \quad \cong -\frac{K-1}{2} \log N, \text{ for very large } N.$$

We have defined S^* only for D^{st} . Now we introduce an extended concept $S^{*\text{I}}$ which applies also to other descriptions, including D^{ind} . [The symbol ' $S^{*\text{I}}$ ' is chosen because the definition applies Method I to be explained later (§6); for an alternative extension $S^{*\text{II}}$ see (10-15).] Let D_i^{ind} be any D^{ind} corresponding to D_i^{st} . Then, according to (2-2)(b), D_i^{ind} has the same value of z as D_i^{st} . Since now $S^*(D_i^{\text{st}})$ depends only on z and Z , we shall take as value of $S^{*\text{I}}(D_i^{\text{ind}})$ the value of $S^*(D_i^{\text{st}})$. We shall do the same for any other D^{ind} with the same z , even if it does not correspond to D_i^{st} , and also for any disjunction of D^{ind} with the same z . Thus we define $S^{*\text{I}}$ as follows:

(3-31) If D_i is either a D^{ind} whose z is z_i or a disjunction of two or more D^{ind} for each of which z has the same value z_i , then

$$S^{*\text{I}}(D_i) =_{\text{df}} \log \frac{z_i}{Z}.$$

Hence:

$$(3-32) \quad \text{For any } D_i^{\text{st}}, \quad S^{*\text{I}}(D_i^{\text{st}}) = S^*(D_i^{\text{st}}).$$

$$(3-33) \quad \text{If } D_i^{\text{ind}} \text{ corresponds to } D_i^{\text{st}}, \quad S^{*\text{I}}(D_i^{\text{ind}}) = S^{*\text{I}}(D_i^{\text{st}}) = S^*(D_i^{\text{st}}).$$

Numerical example. For the sake of simplicity, we take an example with very few cells and round cell numbers. We consider *four statistical descriptions* for the same elements ($N = 10^6$) but with different cell systems. We consider first a D^{st} D_{16} with $K = 16$ cells. Since N is large, we may use (3-28)(b). Let $N_1 = 99,000$, hence $r_1 = 0.099$; $N_2 = 102,000$, hence $r_2 = 0.102$, etc.; the 16 values r_j (not the N_j) are stated in the subsequent table. The example is chosen in such a way that the first eight cell numbers are close to each other and considerably higher than the last eight, which are also close to each other. The values of $r_j \log \frac{1}{r_j}$ are conveniently taken from Dolansky's table.¹

¹ L. and M. P. Dolansky, Table of $\log_2 \frac{1}{p}$, $p \log_2 \frac{1}{p}$, etc. Res. Lab. of Electronics, M.I.T., Techn. Report No. 227, 1952.

From the description D_{16} we derive another description D_8 with 8 cells by merging the first and second cells in D_{16} into a new cell, likewise the third and fourth cells, and so on; thus each r_j -value in D_8 is the sum of two r_j -values in D_{16} . By the same procedure D_4 with $K = 4$ is formed from D_8 and D_2 with $K = 2$ from D_4 .

We see from the table that S^* has nearly the same value for all four descriptions. Each cell in D_2 is divided into two cells of D_4 with nearly equal cell numbers. If the cell numbers were equal, the values of S^* would be equal, because S^* fulfills the second requirement (for large N). Since there is a slight difference between the two cell numbers in each case, a small amount of order is introduced and hence S^* is slightly diminished. The same holds for the transition from D_4 to D_8 , and finally to D_{16} .

The table lists furthermore the minimum and maximum values of S^* for each of the four given values of K , according to (3-29) and (3-30)(a), respectively.

γ (for D_{16})		D_{16} ($K = 16$)	D_8 ($K = 8$)	D_4 ($K = 4$)	D_2 ($K = 2$)
1	0.099	0.3303	0.201	0.4653	
2	0.102	0.3359	0.399	0.5289	
3	0.100	0.3322	0.198	0.4626	
4	0.098	0.3284	0.203	0.4670	0.800 0.2575
5	0.103	0.3378	0.401	0.5286	
6	0.100	0.3322	0.198	0.4626	
7	0.101	0.3341	0.198	0.4626	
8	0.097	0.3265	0.203	0.4670	
9	0.023	0.1252	0.045	0.2013	
10	0.022	0.1211	0.098	0.3284	
11	0.028	0.1444	0.053	0.2246	
12	0.025	0.1330	0.2161	0.102	0.200 0.4644
13	0.026	0.1369	0.050	0.2161	
14	0.024	0.1330	0.052	0.2218	
15	0.025	0.1291	0.052	0.2218	
16	0.027	0.1407			
<hr/>					
	$\sum r_j \log \frac{1}{r_j}$	3.7208	2.7213	1.7218	0.7219
	$\log K$	4.	3.	2.	1.
	$\sum (\dots) - \log K$	-0.2792	-0.2787	-0.2782	-0.2781
	$S^*(D)$	$-2.792 \cdot 10^5$	$-2.787 \cdot 10^5$	$-2.782 \cdot 10^5$	$-2.781 \cdot 10^5$
	S^*_{\min}	$-4 \cdot 10^6$	$-3 \cdot 10^6$	$-2 \cdot 10^6$	-10^6
	S^*_{\max}	-137.38	-67.04	-29.88	-10.29

4

Boltzmann's Entropy Concept

Summary. Boltzmann's concept refers to the state of a gas g at time t consisting of N molecules. The μ -space is divided into cells Q_j of equal volume v^μ . A precise description (D^{prec}) gives the Nn values. The D^{ind} and D^{st} are as before. A macro-description (D^{mac}) states the results of measurements of temperature and pressure in the various parts of g . Boltzmann's definitions of a function H of the cell numbers N_j (4-3) and of entropy (S_B , (4-6)) are given. Both concepts apply to D^{st} . Boltzmann regards all D^{ind} as equally probable (4-8). Then for any D^{st} , S_B is a linear function of the logarithm of its probability (4-10). Boltzmann's definition has certain disadvantages; S_B is dependent upon an arbitrary cell system and is not a continuous function of the values of the magnitudes involved. The abstract concept to be defined later (in Part II) avoids these disadvantages.

Boltzmann¹ considers the state of gas body g at a given time t . g consists of N molecules, each characterized by n magnitudes ϕ_j . This is a special case of the kind explained in §1; but Boltzmann transforms it by the introduction of a cell system into a classification system (as discussed in §§2 and 3). For each magnitude ϕ_j , its interval of admitted values is divided into small intervals of equal length Δ_j . Hereby the μ -space is divided into a system Ω^μ of cells of equal volume $v^\mu = \Delta_1 \dots \Delta_n$. Let K be the number of these cells within the total range R^μ ; then:

$$(4-1) \quad v^\mu = \frac{V^\mu}{K}.$$

These μ -cells Q_j ($j = 1, \dots, K$) are analogous to the cells Q_i in the classification system (§3).

On the basis of a given cell system, Boltzmann makes a distinction, although not clearly and explicitly, which is essentially that between statistical descriptions D^{st} and individual descriptions D^{ind} as explained in §3. He speaks on the one hand of "eventualities" (op. cit. p. 40) or "state distributions" (p. 44), on the other hand of "equipossible cases". It is clear that he means by the former terms D^{st} or states described by D^{st} (called by other authors 'macroscopic states', 'macro-states', 'conditions' or simply 'states') and by the latter term D^{ind} or states described by D^{ind} (called by other authors 'microscopic states', 'micro-states', 'complexions' or

¹ L. Boltzmann, *Vorlesungen über Gastheorie*, Teil I (Leipzig: J. A. Barth, 1896).

simply ‘states’). This interpretation of Boltzmann’s distinction is apparent from the way he explains it by the analogy to drawings from an urn; he shows that here the number of “equipossible cases” “corresponding to an eventuality” with numbers N_1, \dots, N_K is $N!/(N_1! \dots N_K!)$ (p. 39), which is our $z(D_i^{\text{st}})$. For greater clarity, we shall distinguish four kinds of descriptions for g at t .

First, a *precise description* D^{prec} states the nN values of the n magnitudes for the N molecules of g at t . An *individual description* D^{ind} of g at t with respect to Ω^{μ} specifies for each molecule a_i the cell in which its phase point b_i is located, in other words, it specifies for each a_i and each φ_i the interval of the φ_i -scale to which the value $\varphi_i(a_i)$ belongs. A *statistical description* D^{st} for g at t with respect to Ω^{μ} states the cell numbers.

Finally, let us divide the space occupied by (or at least available to) the gas in the vessel (not the phase space) into regions of equal volume (“macro-cells”) just large enough to make it possible to measure in each region the momentary value of temperature and pressure, according to the technological knowledge available at the time in question. A *macro-description* D^{mac} states the results of these measurements.

The logical relations between the four kinds of descriptions are obviously as follows. Once a cell system Ω^{μ} has been chosen, a given D^{prec} , say D_i^{prec} , determines uniquely a certain D_i^{ind} based upon Ω^{μ} such that D_i^{prec} logically imples D_i^{ind} . D_i^{ind} does not imply D_i^{prec} ; it contains less information. The same relation holds between D_i^{ind} and a certain D_i^{st} .

The relation between a D^{mac} and a D^{st} is more complicated. The measurement of a magnitude like temperature T or pressure p is not, strictly speaking, carried out at one time point but during a certain time interval Δt . If the cell numbers N_i were to remain constant during this interval, so that the same D^{st} would hold throughout, then the values of T and p can be determined from this D^{st} with the help of the laws of the kinetic theory of gases. If the cell numbers change during the interval, the same relations hold between the values of T and p on the one hand and the mean values \bar{N}_i ($j = 1, \dots, K$) where \bar{N}_i is the mean of the number of molecules in Q_i during the interval. Let us say that D_i^{st} is *typical* for the interval Δt if the cell numbers stated in D_i^{st} are the mean cell numbers for the interval. If a typical D_i^{st} for an interval Δt is given and a suitable system of macro-cells is chosen, the corresponding D_i^{mac} can be determined (see §5). Inversely, D_i^{mac} determines, with respect to Ω^{μ} , the class C_i^{st} of the corresponding D_i^{st} , that is, those which, taken as typical for Δt , are compatible with D_i^{mac} . The results

of measurements stated by D_i^{mac} are, strictly speaking, not point values but intervals. Therefore the class C_i^{st} contains a very large number of D^{st} . Thus it is a great oversimplification to say, as is often done, that a D^{st} is an appropriate representation of the observed state of a gas.

D_i^{st} , in turn, determines the class C_i^{ind} of the corresponding D^{ind} . D_i^{st} is logically equivalent to the disjunction of these D^{ind} . The same relation holds between D_i^{ind} and the corresponding class C_i^{prec} . (Thus the descriptions may be arranged in trees; see the diagram (10-19) below.) However, while C_i^{st} and C_i^{ind} are finite, C_i^{prec} would be infinite if we meant the D^{prec} to have absolute precision. Although the conception of classical physics, on which our discussions are based, regards it as meaningful to refer not only to finite intervals but also to point values of magnitudes and hence to an absolutely precise "state" as a system of such values, it is impossible to describe in a given language such values or states and it is all the more impossible, not only practically but even theoretically, to know such values or states. Therefore, strictly speaking, we ought to think of a D^{prec} not as something that could actually be given, but rather as an ideal which cannot actually be obtained but can be approached with any precision desired. Any such approximation to a D^{prec} is then really a D^{ind} with the size of the cells determined by the precision obtained.

The introduction of the system of cells is a very ingenious device by which Boltzmann transforms the problem of defining entropy within the framework of the kinetic theory into a problem concerning the simple schema of a K -fold classification. While in the kinetic theory (in its classical form) the possible states constitute an infinite and even nondenumerable set, in the K -schema the number of D^{st} and even that of D^{ind} is finite.

Let f_i be the density in Q_i , i.e., the number of molecules per unit of μ -volume:

$$(4-2) \quad f_i = \frac{N_i}{v^\mu}.$$

Boltzmann defines his H -function for a given D_i^{st} with cell numbers N_i as follows (' Σ ' is short for ' $\sum_{j=1}^K$ '):

$$(4-3) \quad H(D_i^{\text{st}}) = \text{Df} \sum [f_j \ln f_j] v^\mu.$$

Hence

$$(4-4). \text{ (a)} \quad H(D_i^{\text{st}}) = \sum [N_i \ln N_i] - N \ln v^\mu,$$

$$\text{(b)} \quad = -N \left\{ \sum \left[r_i \ln \frac{1}{r_i} \right] - \ln N + \ln v^\mu \right\}.$$

With the approximation (3-9)(b) for large N , (4-1), and (2-1):

$$(4-5) \text{ (a)} \quad H(D_i^{\text{st}}) \cong -\ln z(D_i^{\text{st}}) + N[\ln N - \ln V^\mu + \ln K].$$

$$\text{(b)} \quad \cong -\ln \frac{z(D_i^{\text{st}})}{Z} + N \ln \frac{N}{V^\mu}.$$

Boltzmann defines his statistical concept of entropy, which we denote by ' S_B ', as follows:

$$(4-6) \quad S_B(D_i^{\text{st}}) =_{\text{df}} -k H(D_i^{\text{st}}),$$

where $k = 1.38 \times 10^{-16}$ erg/ $^\circ\text{C}$ (the so-called Boltzmann constant). Hence:

$$(4-7) \quad S_B(D_i^{\text{st}}) \cong k \left\langle \ln \frac{z(D_i^{\text{st}})}{Z} + N \ln \frac{V^\mu}{N} \right\rangle.$$

The purpose of Boltzmann's definition of S_B was to construct on the basis of the kinetic theory a concept corresponding to the *thermodynamic concept of entropy*, S_{th} , which had been introduced by Clausius. It was a stroke of genius for Boltzmann to discover that this purpose could be fulfilled by using a function of the form of his H . What remained was merely to adjust the constant in (4-6) in such a way that the new concept S_B would quantitatively agree with the old concept S_{th} . The latter could be defined in classical physics only relatively, i.e., only the difference of the entropies of two states of a physical system was defined, not their absolute values. Consequently, an arbitrary constant (i.e., a function dependent at most on N , K , and v^μ , but not on the N_i) may be added in (4-3) or (4-4). [This is often done in such a way that H is defined simply by $-N \sum \left(r_i \ln \frac{1}{r_i} \right)$. However, it is to be noted that in this case the additivity (4-11) does not hold.]

The numbers N_i characterize the state of the gas at one time moment; they vary greatly with time, even in very small intervals of time. Therefore H and S_B vary likewise (though at a much lower rate because H is, aside from an additive constant, the weighted mean of $\ln N_i$ over the cells with the N_i as weights, see [4-4](a)).

Since S_{th} is defined in terms of measurable magnitudes, it can at best agree with a mean of S_B for a time interval Δt (see the above remarks on D^{mac}). We shall come back to this point in the next section.

Boltzmann introduces a probability distribution (denoted here by ' pr_B ') suggested by the experimental results, though not uniquely determined by them. (The term 'probability' is here to be understood in the statistical, not the inductive or logical sense.²) He first regards the probability that the phase point of any given molecule belongs to a given sub-region of the range R^{μ} as proportional to the volume of the sub-region. Therefore the probability that the phase point belongs to a given cell is $1/K$. Secondly, he takes the probabilities for different molecules as independent of each other. Therefore the $Z D^{ind}$ have equal probabilities:

$$(4-8) \quad \text{For every } D_i^{ind}, \ pr_B = \frac{1}{Z}.$$

A D^{st} is a disjunction of the z corresponding D^{ind} ; therefore:

$$(4-9) \quad pr_B(D_i^{st}) = \frac{z(D_i^{st})}{Z}.$$

Hence with (4-7):

$$(4-10) \quad S_B(D_i^{st}) \cong k \left[\ln pr_B(D_i^{st}) + N \ln \frac{V^{\mu}}{N} \right].$$

Thus $S_B(D_i^{st})$ is a linear function of the logarithm of the probability of D_i^{st} . Boltzmann emphasizes this close connection between entropy and probability (pp. 38 ff); he shows that the second law of thermodynamics (see below, §5) is derivable as a probability statement (pp. 42, 60).

(4-11) *Additivity of H and S_B .* If the gas g at t consists of p spatial parts g_m ($m = 1, \dots, p$), then, with respect to D^{st} for g and for its parts,

$$(a) \quad H(g) = \sum_{m=1}^p H(g_m),$$

$$(b) \quad S_B(g) = \sum_{m=1}^p S_B(g_m).$$

² Concerning this distinction, see (Prob.), §12.

Proof. The parts g_m correspond to non-overlapping regions R_m of cells in the μ -space. Let $N^{(m)}$ be the number of molecules in g_m ;

$\sum_m N^{(m)} = N$. Then from (4-4)(a):

$$H(g_m) = \sum_{Q_i \text{ in } R_m} [N_i \ln N_i] - N^{(m)} \ln v^\mu.$$

Hence $\sum_m H(g_m) = \sum_{Q_i \text{ in } R^\mu} [N_i \ln N_i] - N \ln v^\mu = H(g)$. This is (a). (b) follows from (a) with (4-6).

We have, as is customary, assumed that the cells Q_i in the μ -space have equal volumes v^μ . This is essential for the derivation of (4-4), (4-5), (4-9), and (4-10). It may be remarked that for a definition of the form (4-3) this assumption is not necessary. We could use cells Q_i with different volumes v_i and define H by $\sum (f_i \ln f_i) v_i$. Suppose that two cells Q'_i and Q''_i with equal densities $f' = f''$ are merged into one new cell Q_i with $v_i = v'_i + v''_i$ (v'_i and v''_i need not be equal) and $N_i = N'_i + N''_i$. Then in Q_i , $f = f' = f''$. The two terms $(f' \ln f')v'_i$ and $(f'' \ln f'')v''_i$ in the \sum for the first system are replaced for the second system by one term $(f \ln f)v_i$, which is equal to the sum of the two terms. Thus \sum remains unchanged. These considerations will be of importance for our construction of an abstract concept of entropy (in Part II). There we shall make use of a system of cells (called "environments") of different volumes.

In the present schema, with equal cell sizes v^μ , we can make a fusion or a subdivision only if we do it for all cells simultaneously. A consideration like the one above leads to the following result.

(4-12) Invariance of H and S_B with respect to a subdivision of calls. Let D be a D^{st} for N molecules, based on a system Ω^μ of K cells Q_j ($j = 1, \dots, K$) of volume v^μ with cell numbers N_j . Let D' be a D^{st} based on the system Ω'^μ of pK cells of volume v'^μ/p ; each previous cell Q_j is here subdivided into p cells with equal cell numbers $N_{j1} = N_{j2} = \dots = N_{jp} = N_j/p$. Then

$$(a) \quad H(D) = H(D')$$

$$(b) \quad S_B(D) = S_B(D').$$

Proof. In D' , for every i ($= 1, \dots, p$), $r_{ji} = N_{ji}/N = N_j/pN = r_j p$. In

(4-4)(b), the term $r_j \ln \frac{1}{r_j}$ in \sum for D is therefore replaced in \sum' for D'

by p equal terms $(r_i/p) \ln(p/r_i)$; their sum is $r_i[\ln(1/r_i) + \ln p]$. Thus, since $\sum r_i = 1$, $\sum' = \sum + \ln p$. $\ln v^\mu$ is replaced by $\ln(v^\mu/p) = \ln c^\mu - \ln p$. Therefore $H(D') = H(D)$.

The following theorems state maxima and minima for given N and V^μ (or N , K , and v^μ).

(4-13) For a uniform distribution (every $N_i = N/K$, every $r_i = 1/K$):

(a) $\sum\left(r_i \ln \frac{1}{r_i}\right)$ has its maximum $\ln K$.

(b) H has its minimum $-N \ln \frac{Kv^\mu}{N} = -N \ln \frac{V^\mu}{N}$ (from (4-1)).

(c) S_B has its maximum $kN \ln \frac{V^\mu}{N}$.

(4-14) For the case of all molecules concentrated in one cell (one $N_i = N$, one $r_i = 1$, all others 0):

(a) $\sum\left(r_i \ln \frac{1}{r_i}\right)$ has its minimum 0.

(b) H has its maximum $N(\ln N - \ln v^\mu) = N \ln \frac{NK}{V^\mu}$.

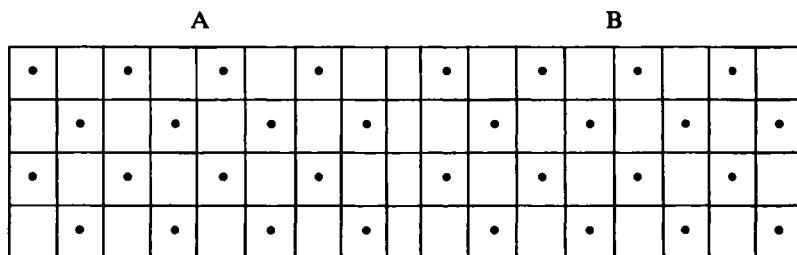
(c) S_B has its minimum $-kN(\ln N - \ln v^\mu) = -kN \ln \frac{NK}{V^\mu}$.

The form proposed by Boltzmann for a definition of entropy within the kinetic theory of gases has, in spite of its great merits, also certain *disadvantages*, even from the point of view of classical physics. They have been pointed out repeatedly. We shall discuss two of them in greater detail, because they are relevant for our later discussions in Part II. The first disadvantage consists in the fact that S_B depends upon a cell system. This cell system is not determined by the state of g at t or by its D^{prec} , i.e. the values u_{ij} of the magnitudes ϕ_i for the elements a_i ; it is rather chosen arbitrarily. Even for the same number K of cells, different locations and different directions of the cells may be chosen. A different choice, even with the same K , leads in general to a different value of S_B for the same given D^{prec} . And the choice of K , which determines also the size $v^\mu = V^\mu/K$ of every cell, presents an insoluble dilemma. If we choose a small K and hence large cells, the determination of S_B will be imprecise, because any differences in the density of phase points *within* the cells are neglected. If we subdivide the cells and thus take $K' > K$, we obtain in general another (greater) value of S_B , which is more precise (see the example (3-34) for S^* .) It might seem not implausible to assume that, at least theoretically on the basis of

the classical conception, it would be possible to increase at will the precision of the determination of S for a given D^{prec} by taking a sufficiently high K , in other words sufficiently small cells. One might perhaps even think that S could be defined theoretically, though not obtained actually, by the limit for $v^\mu \rightarrow 0$. [Perhaps Boltzmann himself thought so. At one place (pp. 45 f.) he tries to justify his procedure of writing differentials ' $d\xi$ ' etc. instead of finite intervals of the coordinates and volume differentials in the phase space instead of finite volumes for the cells. However, his reasoning at this point is not very clear; and at the end of the discussion he says explicitly, that the use of differential equations in an atomistic theory can only be regarded as an approximation.] However, first of all there is no such convergence; the numerical density $f_i = N_i/v^\mu$ with decreasing v^μ becomes either 0 or grows beyond any bound; there is no positive finite limit. Secondly, a cell system with a too large K and hence very small though still positive v^μ , is unsuitable for Boltzmann's definition of S_B . In such a system, differences in the degree of disorder and hence differences in the thermodynamical entropy S_{th} are no longer represented by corresponding differences in the cell numbers N_i and in S_B . This becomes clear by the following examples. For simplicity, we disregard the momentum components and therefore identify the cells with spatial cells in the container of the gas. We consider two gas states G_1 and G_2 (see the examples in figs. 1 and 2). The vessel containing the gas is divided by a wall into two parts A and B . In G_1 all molecules are uniformly distributed in part A . Therefore there is a considerable difference between the densities in the two parts. This difference could be exploited to gain mechanical energy from the gas. It is clear that S_{th} in a situation of this kind is low. On the other hand, in G_2 there is uniform density throughout the container and hence high S_{th} . We form two D^{st} for G_1 , D_1 and D'_1 , with two different cell systems. D_1 uses $K = 2N$ small cells. Each of the N cells in part A contains one molecule. D'_1 uses $K' = K/m = 2N/m$ large cells (where m is an even number and N is divisible by m), each cell comprising m of the cells in D_1 . (In the diagrams,

A								B							
•	•	•	•	•	•	•	•								
•	•	•	•	•	•	•	•								
•	•	•	•	•	•	•	•								
•	•	•	•	•	•	•	•								

Figure 1. Gas state G_1 : all molecules in Part A.

Figure. 2. Gas state G_2 : uniform distribution.

we take for simplicity very small numbers: $N=32$, $m=4$; the $K=64$ small cells are represented by the small squares; the $K'=16$ large cells are indicated by heavy lines.) Let the size of the small cells be $v^u = 1$; then that of the large cells is $v^{u'} = m$. We obtain by

$$(4-4)(a) H(D_1) = 0, \quad H(D'_1) = \frac{N}{m} \cdot m \ln m - N \ln m = 0. \quad G_2 \text{ is likewise}$$

described by two D^s , D_2 and D'_2 with the same cell systems as D_1 and D'_1 , respectively. In D_2 , just as in D_1 , all cell numbers are 1 or 0; therefore $H(D_2) = H(D_1)$ and, with (4-6), $S_B(D_2) = S_B(D_1) = 0$. Thus D_1 and D_2 have equal values of S_B in spite of the great difference in the thermodynamic situation between G_1 and G_2 . On the other hand, in D'_2 , all cell numbers are $m/2$. Therefore $H(D'_2) = \frac{2N}{m} \cdot \frac{m}{2} \ln \frac{m}{2} - N \ln m = N \ln 2$. Hence $S_B(D'_2) = kN \ln 2 > S_B(D'_1)$.

Thus D'_1 and D'_2 reflect the difference in S_{th} because they use a system of larger cells.

Generally speaking, K should be chosen sufficiently small to make the mean cell number N/K considerably greater than 1; in other words, K should be considerably smaller than N . (In our example, in D_1 and D_2 $N/K = 1/2$, and in D'_1 and D'_2 $N/K = 2$, which is actually still too small.)

The second disadvantage of Boltzmann's definition is perhaps even more serious from the point of view of classical physics. We assume now that for a given D^{prec} , a fixed cell system with a suitable K has been chosen. We consider a continuous transition from the given D^{prec} into a slightly different one, in other words, a continuous variation of some of the values u_{ij} (or of just one of them, say u_{11}). Then S_B may change in a discontinuous way; in general there will be a jump in the value of S_B whenever a phase point crosses the boundary between two adjacent cells. Thus S_B is not a continuous function of the values u_{ij} .

It seems desirable to construct a statistical concept S in such a

way that the two disadvantageous features of S_B are avoided, in other words, such that the following two requirements are fulfilled:

(4-15) First requirement. S should not be dependent upon the arbitrary choice of a cell system.

(4-16) Second requirement. $S(D^{\text{prec}})$ should be a continuous function of the values u_{ij} in D^{prec} . The abstract concept S^{**} to be defined in Part II will fulfill these two requirements.

5

Boltzmann's H-theorem

Summary. Boltzmann stated his H-theorem as a statistical counterpart to the second law of thermodynamics. It is based on certain assumptions which are plausible but not strictly proved. Boltzmann's S_B is in agreement with thermodynamic entropy S_{th} under the condition of (approximate, local) equilibrium.

The second law of thermodynamics, the entropy principle, says that in an isolated system, e.g., a gas body enclosed in adiabatic walls, the thermodynamic entropy S_{th} always increases. (The entropy may remain constant if the process is reversible; this, however, is only an ideal case which can be approximated but not exactly realized.) If g is an isolated system, the total energy E remains constant:

$$(5-1) \quad E(g, t_1) = E(g, t_2).$$

As a statistical counterpart to the second law, Boltzmann derived the following result by a detailed analysis of the changes in the value of H due to collisions of molecules and other processes.

(5-2) *Boltzmann's H-theorem.* Let D_1^{st} and D_2^{st} be two S^{st} describing the states of an isolated gas body g with constant energy E at the times t_1 and $t_2 (> t_1)$ respectively. Suppose that $H(D_1^{st})$ is considerably higher than the mean of H for the given E . Then with overwhelming probability $H(D_2^{st}) < H(D_1^{st})$, and hence $S_B(D_2^{st}) > S_B(D_1^{st})$.

This means that in the course of time H will with great probability decrease until it is near its mean, which is very close to its minimum. Hence S_B will with great probability increase until it is near its mean, which is very close to its maximum (for the given E).

Boltzmann's derivation of his theorem was supplemented and refined by later authors. However, it is now generally recognized that the derivation does not constitute a rigorous proof. It consists rather in showing (1) that it seems plausible to assume that certain factors in the situation cancel each other out on the average for a large number of cases, and (2) that if they do, H decreases. We shall later (in §9) make some remarks about the unproved assumptions involved.

Suppose that for an isolated body of gas the total energy E , the total mass M , and the number of molecules N are known. Then the

mass m of each molecule is $m = M/N$. For any molecule in a given cell Q_j , its momentum and hence, with m , its kinetic energy ε_j is known. We consider the simplest case, that of a diluted monatomic gas; here E is the sum of the kinetic energies of the molecules. Therefore only those D^{st} (with the given N and K) are possible which fulfill the following “energy condition”:

$$(5-3) \quad \sum_i N_i \varepsilon_i = E.$$

Among these D^{st} , let D' with cell numbers N'_i be the one for which S_B has its maximum (or its mean, which makes practically no difference) for the given E . Then the N'_i show the Maxwell-Boltzmann distribution:

$$(5-4) \quad N'_i = e^{-\alpha - \beta \varepsilon_i},$$

where the parameters α and β are uniquely determined by $\sum N_i = N$ and (5-3). In this case the temperature T and the pressure p have the same value throughout the gas; hence the gas is in *thermodynamic equilibrium*. Therefore the distribution (5-4) of the N_i is said to characterize *statistical equilibrium*. The values of p and T are connected with the parameter β as follows:

$$(5-5) \quad p = \frac{N}{\beta V},$$

where V is the volume of the gas body. Hence, with the gas law

$$(5-6) \quad pV = NkT, \text{ where } k \text{ is Boltzmann's constant, we obtain:}$$

$$(5-7) \quad T = \frac{1}{k\beta}.$$

In the case of a diluted monatomic gas, we have

$$(5-8) \quad E = \frac{3}{2} NkT.$$

Thus the distribution (5-4) can also be expressed in the following forms, with ‘ C ’ = ‘ $e^{-\alpha}$ ’:

$$(5-9) \quad (a) \quad N'_i = Ce^{-\frac{\varepsilon_i}{kT}},$$

$$(b) \quad = Ce^{-\frac{3N}{2kT} \varepsilon_i}.$$

Now we can formulate more precisely the conditions for an agreement between Boltzmann's S_B and thermodynamic entropy S_{th} . First, S_{th} is defined only for a state of the gas in which there is at least approximate thermodynamic equilibrium either in the whole or at least locally, i.e., in each spatial part. If the gas is in such a state, the numbers N_i and hence S_B still undergo rapid fluctuations. Therefore the agreement cannot possibly exist at all moments. The second condition is that we take the description D' with the numbers N'_i given by (5-4) or (5-9). This is a typical D^{st} (in the sense explained in §4); its S_B -value is the mean for any sufficiently long time interval. The agreement holds for this S_B , which is given by:

$$(5-10) \quad S_B(D') = -k[\sum (N'_i \ln N'_i) - N \ln v^u].$$

S_{th} is defined in classical physics only up to an arbitrary additive constant. Under the conditions mentioned, S_B and S_{th} are numerically equal if the constant is suitably chosen.

Suppose that the vessel containing the gas g is divided by adiabatic walls into several parts. Suppose further that, at the given time, for each part g_n of g the two conditions are fulfilled: that is to say, g_n is in or near to thermodynamic equilibrium and we take a corresponding D^{st} for g_n which is nearly in statistical equilibrium. Then the agreement between S_B and S_{th} holds for each part g_n . Now the value of S_{th} for the whole gas body is equal to the sum of its values for the parts, and S_B is likewise additive (4-11). Therefore the agreement holds also for the whole.

In order to obtain a model for a process in which the entropy increases toward its maximum, we again consider the vessel divided into parts, with thermodynamic equilibrium in each part; T and p may have different values, high in some parts, lower in others. But now we perforate each wall by a hole so small that the equalization of T and p takes place very slowly. Then at any time t during this process, there is still approximate thermodynamic equilibrium in each part. And if again we take at any t the typical D^{st} , i.e., the combination of typical D^{st} for the parts, we have still agreement between S_B and S_{th} . According to the second law, S_{th} increases during the process until it reaches its maximum. And according to Boltzmann's H -theorem, S_B does likewise.

6

The Problem of the Definition of Entropy for other than Statistical Descriptions

Summary. With this section, the critical examination of the statistical concept of entropy begins. Boltzmann applied his concept S_B only to D^{st} . Later physicists extended its use to other descriptions, including D^{ind} . We discuss two basically different methods for this extension, which we call Methods I and II. First, we state two principles which seem to be generally accepted (tacitly) for any purely physical magnitude M (as distinguished from other, e.g., logical or epistemological concepts): (1) If M is a function of certain parameters, and two descriptions state the same values for these parameters, then we must ascribe to M the same value on the basis of either description (6-1); (2) if M has the same value on the basis of each of several descriptions, then it has also the same value on the basis of their disjunction (6-2). Method I, which I advocate, defines the extended concept S_B^I in such a way that the two principles are fulfilled; therefore S_B^I is a purely physical concept. For a D^{st} , S_B^I has the same value as S_B . For a D^{prec} or a D^{ind} , S_B^I has the same value as for the corresponding D^{st} . If each of several D^{st} has the same value of S_B , then this is also the value of S_B^I for their disjunction. It is shown that S_B^I for any of these descriptions is in accord with S_{th} . In Method II, an extended concept S_B^{II} is defined for any description D_i in such a way that its value is, aside from an additive constant, proportional to the logarithm of the probability of D_i (where, in accordance with Boltzmann's conception, all D^{ind} are taken as equally probable). It seems that this concept S_B^{II} is customarily accepted by physicists. However, it is shown that S_B^{II} , in contrast to S_B^I , is not generally in accord with thermodynamic entropy except for a D^{st} . It violates the two principles and therefore is not a purely physical concept; it is even a purely logical concept. Therefore it is entirely out of line with Boltzmann's aim to construct a physical, statistical concept corresponding to the thermodynamic concept of entropy.

Boltzmann's concept of entropy S_B was originally applied by himself and other physicists only to statistical descriptions (or the states described by them). The way in which Boltzmann and others talk about "the state" of a gas g at a given time t as described by a D^{st} suggests that perhaps a D^{st} more than a D^{ind} appears to them as representing "the objective physical reality." However, on the basis of the classical conception, the objective state in its quantitative aspects is the set of the actual values of the magnitudes ϕ_i ; hence its

representation is a D^{prec} rather than a D^{st} . Once a cell system Ω is chosen, a given D_i^{prec} determines uniquely the corresponding D_i^{ind} and the corresponding D_i^{st} . While D_i^{ind} is still an approximation of D_i^{prec} , D_i^{st} is a very weak consequence of D_i^{ind} and therefore a very limited description of the state, although very useful for certain purposes.

After the time of Boltzmann, physicists applied the concept S_B also to other kinds of descriptions, including D^{ind} . But the question whether and how this extension is to be made seems still controversial. Or rather, there seems to be a lack of explicit agreement without controversy. At least, I have not seen so far any discussion in print concerning the possible modes of extension. Usually the extension is made tacitly, without an explicit formulation of a definition; still less are reasons given for the choice of the method.

We shall explain two quite different methods of extension for S_B . But first a brief remark about the view that S_B must be restricted to D^{st} and that any extension of it and, in particular, any application to D^{ind} is illegitimate. (I have not found this view in print, but only in conversations.) I find it hard to understand on what reasons this view might be based; it seems to me to give to the concept S_B a peculiar role by permitting it to violate the subsequent principle (6-1)(A). On the basis of the conception of classical physics, this principle seems to be tacitly assumed for any ordinary, purely physical measurable magnitude M .

(6-1) Principle of physical magnitudes. Let M be a physical magnitude applicable to the state of a physical system s at time t with respect to a reference system R . Let M be defined in terms of n other magnitudes m_1, \dots, m_n ; i.e., if the values of these magnitudes are given (with a certain precision) then the value of M is thereby uniquely determined (with a certain precision). Let D_1 and D_2 be two descriptions of s at a specified time t_0 with respect to R such that D_1 and D_2 give the same values of $m_1(s, t_0), \dots, m_n(s, t_0)$. If M is defined in such a way that, on the basis of D_1 , a certain value x is ascribed to M for s at t_0 with respect to R , then:

- (A) on the basis of D_2 , a value should be ascribed to M ,
- (B) this value must be equal to x .

Part (A) of the principle is not satisfied by the conception of S_B mentioned above. Part (B) will be used later as an argument for my Method I as against the customary Method II.

Example. Suppose that s is a straight rigid bar, which at time t_0 is at rest in the spatial coordinate system R . D_1 gives the six spatial

coordinates of the two end points of the bar at time t_0 with respect to R . D_2 gives the same values of the coordinates, but adds that s is made of copper and that its temperature at t_0 is 100°C. Let M be the length of s at t_0 with respect to R . Here it seems obvious that on the basis of D_2 the same value must be ascribed to M as on D_1 . But now suppose that M' is a new magnitude proposed by a physicist. He defines it in such a way that its value for s at t with respect to R on the basis of descriptions like D_1 is determined by a function f of the six coordinates mentioned; hence on D_1 , the value of M' is equal to the value of f for the coordinate values stated in D_1 . Although the author of a new concept is certainly free to decide how to define it, within the context of classical physics most physicists would find it strange if the author declared that on D_2 no value should be ascribed to M' . And if he constructed the definition in such a way that M' had a value on D_2 , but not the same as the value on D_1 , then many physicists would say that M' is not a purely physical concept; it might be a logical concept dependent on the logical properties of the description D in question, or an epistemological concept dependent on the state of knowledge of an observer as formulated in D . Or again, it might be a concept of a mixed nature, dependent on both physical and non-physical, say logical or epistemological, data.

These considerations suggest that the decision concerning the application of S_B to descriptions of other kinds will depend upon the general nature and purpose of the statistical concept of entropy.

The concept of entropy in thermodynamics (S_{th}) had the same general character as the other concepts in the same field, e.g., temperature, heat, energy, pressure, etc. It served, just like these other concepts, for the quantitative characterization of some objective property of a state of a physical system, say, the gas g in the container in the laboratory at the time t . When the kinetic theory of gases was developed, one of the essential problems was to find counterparts to the thermodynamic concepts within this theory, to be defined in terms of statistical mechanics. Thus Boltzmann's problem was not essentially different from the analogous problems for the other concepts. As, for example, the statistical equivalent of temperature was found to be proportional to the mean kinetic energy of the molecules, Boltzmann found that the statistical equivalent of entropy was a certain logarithmic function of the cell numbers. It is true that the situation with respect to entropy differs in one respect from that of the other concepts: Boltzmann introduced as a new constituent the cell system, which is not determined by the objective state of the physical system itself, but must be

chosen arbitrarily. However, this fact, although not unimportant, does not mark a fundamental difference in the general character of the concept. First, if the number of cells is made sufficiently large, then under ordinary conditions as in a gas the influence of a change of the cell system is numerically small. And secondly, it is possible to modify the Boltzmann concept in such a way that the arbitrary cell system disappears; this will be shown later (in Part II). Therefore it seems to me that the statistical concept of entropy introduced by Boltzmann was meant by him as a purely physical concept just like the original thermodynamic concept of entropy.

The following principle (6-2) seems likewise to be generally accepted by tacit convention for any purely physical concept.

(6-2) Principle of disjunction. Let M be a physical magnitude. Suppose that, for a given physical system at a given time with respect to a given reference system R , M has the same value on the basis of each of several descriptions. Then H has the same value also on the basis of the disjunction of these descriptions.

Example. Let M be the length of a bar s at time t_0 in R , as in the previous example. Let s be parallel to the x -axis. Let the x -coordinates of the endpoints of s at t_0 be stated in D_1 as 1.3 and 1.5, in D_2 as 2.3 and 2.5, and in D_3 as 3.3 and 3.5. Thus $M(D_1) = M(D_2) = M(D_3) = 0.2$. The principle of disjunction says that on the basis of the disjunction D_d of the three descriptions, M is likewise 0.2.

Assume again that M' is a new magnitude proposed by a physicist. Suppose that his definition of M' is such that, on the basis of each of the descriptions D_1 , D_2 , and D_3 , $M' = 0.2$. He is certainly free to construct his definition as he sees fit, as long as it is consistent. Suppose that definition is such that, on the basis of the disjunction D_d , it leads to a value of M' different from 0.2. We cannot say that the definition is inconsistent; but again, as in the previous example, many physicists would say that M' is then not a purely physical concept.

Both principles seem rather obvious. They may even appear to the reader as too obvious and trivial to bother about their explicit formulation. However, we shall find that the customary form of the concept of entropy in statistical mechanics violates both principles.

Now we shall explain the alternative Methods I and II for extending the definitions of H and S_B . We denote the concepts resulting from Method I by ' H^I ' and ' S_B^I ' and those from Method II by ' H^{II} ' and ' S_B^{II} '.

The main characteristic of *Method I*, which I advocate, is the conception of H and S_B as purely physical concepts. Consequently, the concepts H^I and S_B^I will be defined in such a way that they satisfy the principles (6-1) and (6-2). Let a D^{st} for g at t be given, say D_1^{st} . It is based on a space-time coordinate system and a system Ω^u of K cells in the μ -space; these two systems together constitute the reference system R in the principle (6-1). $H(D^{st})$ is a function of the cell numbers N_1, \dots, N_K ($N = \sum N_i$, K is given by R , k is a universal constant with respect to R). Thus these cell numbers are in this case the magnitudes m_1 , etc. referred to in the principle. Let D_2 be another description of any kind for g at t with respect to R which states the same cell numbers N_i as D_1^{st} . Then, according to the principle (6-1), the definition of H is to be extended in such a way that

$$(6-3) \quad H^I(D_2) = H(D_1^{st}).$$

We define S_B^I in analogy to (4-6), but now for any description D_i to which H^I is applicable:

$$(6-4) \quad S_B^I(D_i) =_{Df} -kH^I(D_i).$$

Then S_B^I will likewise satisfy (6-1).

In order to fulfil (6-3), we define

$$(6-5) \quad H^I(D_i^{ind}) =_{Df} H(D_i^{st}), \text{ where } D_i^{st} \text{ is the } D^{st} \text{ corresponding to } D_i^{ind}.$$

Hence with (6-4):

$$(6-6) \quad S_B^I(D_i^{ind}) = -kH(D_i^{st}) = S_B(D_i^{st}).$$

Let D_i^{d-ind} be a disjunction of several D^{ind} which correspond to the same D_i^{st} and therefore have the same value of S_B^I . In order to fulfil the principle of disjunction, we assign to D_i^{d-ind} likewise this value of S_B^I :

$$(6-7) \quad S_B^I(D_i^{d-ind}) =_{Df} S_B(D_i^{st}).$$

Let D_i^{d-st} be a disjunction of several D^{st} , including D_i^{st} , which have the same value of S_B as D_i^{st} . Again, in order to fulfil the principle of disjunction, we assign the same value also to the disjunction:

$$(6-8) \quad S_B^I(D_i^{d-st}) =_{Df} S_B(D_i^{st}).$$

We can extend the definition of H^I also to a D^{prec} in accordance with (6-3). In this case, however, we must insert into the definiendum a reference to a cell system Ω^μ , because a D^{prec} , in contradistinction to a D^{st} or a D^{ind} , does itself not refer to a cell system:

(6-9) $H^I(D_i^{prec}, \Omega^\mu) =_{Df} H(D_i^{st})$, where Ω^μ is any cell system and D_i^{st} is that D^{st} based on Ω^μ which corresponds to D_i^{prec} (i.e., is logically implied by it). Hence with (6-4):

$$(6-10) \quad S_B^I(D_i^{prec}, \Omega^\mu) = S_B(D_i^{st}).$$

Let us examine the relation between S_B^I and thermodynamic entropy S_{th} . Suppose that g at time t_1 is in a state of near-equilibrium and that D_i^{st} is a typical description of g at t_1 . Then $S_B^I(D_1^{st}) = S_B(D_1^{st}) = S_{th}(g, t_1)$. Then, for any D^{ind} corresponding to D_1^{st} , for any disjunction of several D^{ind} of this kind, and for any D^{prec} corresponding to D_1^{st} (with respect to the cell system used in D_1^{st}), the value of S_B^I is equal to $S_B(D_1^{st})$ and hence to $S_{th}(g, t_1)$. Furthermore, let D_i be a disjunction of several D^{st} , including D_1^{st} , each of which has the same S_B -value as D_1^{st} . Then $S_B^I(D_i) = S_B(D_1^{st}) = S_{th}(g, t_1)$. Thus *the value of S_B^I for any of these descriptions is in accord with S_{th} .*

Our earlier extension of the concept S^* (for a simple classification, cf. §3) to D^{ind} (see (3-31)) was likewise in accordance with Method I.

Method II, which seems to be most frequently used when an extension of S_B is made proceeds as follows. For a given system of N molecules and a given system Ω^μ of K cells, let D_i be a disjunction of m distinct D^{ind} ($1 \leq m \leq Z = K^N$; the cell numbers need not be the same in the several components of D_i). Method II is characterized by defining $H^{II}(D_i)$ by a function of the form $-\ln m + \text{const.}$, and hence $S_B^{II}(D_i)$ by $k(\ln m - \text{const.})$. The constant is chosen in various ways, sometimes as 0. Let us take it in such a form that S_B^{II} is in accordance with S_B in the case of a D^{st} . Therefore we define:

$$(6-11) \quad \begin{aligned} \text{(a)} \quad S_B^{II}(D_i) &=_{Df} k \left\{ \ln m - N \ln \frac{NK}{V^\mu} \right\} \\ \text{(b)} \quad &= k(\ln m - N \ln N + N \ln V^\mu]. \end{aligned}$$

Any D_i^{st} is logically equivalent to the disjunction of the z corresponding D^{ind} .

Therefore

$$(6-12) \quad S_B^{\text{II}}(D_i^{\text{st}}) = k \left\{ \ln z(D_i^{\text{st}}) - N \ln \frac{NK}{V^\mu} \right\}.$$

As we see from (4-5)(a) and (4-6), this is indeed equal to $S_B(D_i^{\text{st}})$ (with the approximation (3-9)(b), which is always used in statistical mechanics); hence S_B^{II} is indeed an extension of Boltzmann's concept S_B .

For a D^{ind} , $m = 1$. Therefore:

$$(6-13) \quad S_B^{\text{II}}(D_i^{\text{ind}}) = -kN \ln \frac{NK}{V^\mu}.$$

For the concept S^* (§3), we defined an extension $S^{*\text{I}}$ in (3-31), which applies Method I. A definition of $S^{*\text{II}}$, which applies Method II, will be given later in (10-15).

For Boltzmann's probability distribution pr_B (4-8), we have:

$$(6-14) \quad pr_B(D_i) = \frac{m}{Z},$$

where $Z = K^N$. Therefore $\ln pr_B(D_i) = \ln m - N \ln K$. Hence

$$(6-15) \quad S_B^{\text{II}}(D_i) = k \left\{ \ln pr_B(D_i) - N \ln \frac{N}{V^\mu} \right\}.$$

Sometimes S_B^{II} is directly defined in terms of probability in the form $k(\ln pr_B(D_i) - \text{const.})$, e.g., by (6-15) or with const. = 0. We see that the simple relation between S_B and pr_B (4-10), which for our form S_B^{I} holds only in the case of D^{st} , holds for S_B^{II} generally with any D_i of the kind described, including all D^{ind} . This result—and another related result concerning the amount of information (§10)—may at first glance appear as an advantage of Method II in comparison with I; and it was perhaps the main motivation for those who accepted Method II.

However, Method II has also very serious disadvantages. We shall show that S_B^{II} is in most cases not in agreement with thermodynamic entropy S_{th} . Moreover, S_B^{II} is not a purely physical concept but a

purely logical concept. This will become clear by the consideration of some examples.

The first two examples use a simple case of the model explained in §5. The vessel containing the gas g is divided by a wall into two parts A and B . During a time interval around t_1 long enough for repeated measurements, it is found that the temperature throughout part A remains T_A , and throughout B T_B ; similarly the pressure throughout A remains p_A , and throughout B p_B , where $T_A > T_B > 0$ and $p_A > p_B > 0$. Later the wall is removed and a process of equalization of T and p takes place. During a still later interval around t_2 , T is found to have the value T_2 and p the value p_2 throughout the vessel. Then, according to the second law:

$$(6-16) \quad S_{\text{th}}(g, t_2) > S_{\text{th}}(g, t_1).$$

Let D_A^{st} be a typical D^{st} for part A and D_B^{st} for part B for the first time period. Let D_1^{st} be the combination of D_A^{st} and D_B^{st} ; thus D_1^{st} is a typical D^{st} for the whole gas for the first period. Let D_2^{st} be typical for the second period. Then we see from the discussion in the preceding section:

$$(6-17) \quad (\text{a}) \quad S_B(D_1^{\text{st}}) = S_{\text{th}}(g, t_1),$$

$$\quad (\text{b}) \quad S_B(D_2^{\text{st}}) = S_{\text{th}}(g, t_2);$$

and

$$(6-18) \quad S_B(D_2^{\text{st}}) > S_B(D_1^{\text{st}}),$$

in accordance with Boltzmann's H-theorem.

Example 1. We have assumed D_1^{st} and D_2^{st} to be true (typical descriptions of g at t_1 and t_2 , respectively. Now let D_1^{ind} be the one true D^{ind} corresponding to D_1^{st} , and likewise D_2^{ind} for D_2^{st} . (Although no observer can know which of the D^{ind} is the true one, we can, within the conception of classical physics, refer to it hypothetically, just as we can refer to the unknown value u_1 of a magnitude u at time t_1 .) If $z(D_i^{\text{st}})$ were = 1, then all molecules would be concentrated in a very small part of the vessel. Since this is not the case in D_1^{st} , we have $z(D_1^{\text{st}}) > 1$. Therefore, with (6-13) and (6-12):

$$(6-19) \quad S_B^{\text{II}}(D_2^{\text{ind}}) = S_B^{\text{II}}(D_1^{\text{ind}}) < S_B(D_1^{\text{st}}).$$

Hence with (6-17)(a) and (6-16):

$$(6-20) \quad S_B^{II}(D_2^{\text{ind}}) = S_B^{II}(D_1^{\text{ind}}) < S_{\text{th}}(g, t_1) < S_{\text{th}}(g, t_2).$$

Thus we see that S_B^{II} has the same value for D_1^{ind} and D_2^{ind} , a value that is smaller than S_{th} for t_1 and still smaller than S_{th} for t_2 . This shows that S_B^{II} is not in accord with S_{th} . D_1^{ind} is not only a true description of the state of the gas at t_1 but a description from which D_1^{st} is deducible, which in turn yields the value $S_B(D_1^{\text{st}})$ which agrees with $S_{\text{th}}(g, t_1)$; the situation for D_2^{ind} is analogous. Thus we see that, although D_1^{ind} and D_2^{ind} contain the relevant information for the determination of correct S -values (i.e., those which agree with S_{th}), S_B^{II} assigns to them other values. More generally, we see from (6-13) that S_B^{II} has the same value for all D^{ind} (for given V^u , K , and N), although these D^{ind} describe all possible states of g with a great variation in S_B , in S_{th} , and hence in the amount of available energy which could be transformed into mechanical work. *Thus S_B^{II} for D^{ind} has no connection whatever with thermodynamic entropy and available energy.* Since $S_B^{II}(D_1^{\text{ind}}) < S_B^{II}(D_1^{\text{st}})$, S_B^{II} violates the principle (6-1)(B). *Therefore S_B^{II} is not a purely physical concept.*

Example 2. We use again D_1^{st} and D_2^{st} as described above. It follows from (6-18) that $z(D_2^{\text{st}}) > z(D_1^{\text{st}})$. Let us suppose now that the differences $T_A - T_B$ and $p_A - p_B$ are small (one of them may even be 0) but at least one of them is still noticeable, so that the results of the measurements still lead to (6-16). Then $z(D_2^{\text{st}})$ is not very much greater than $z(D_1^{\text{st}})$. Let n be a positive integer (say the smallest one) such that

$$(6-21) \quad n \times z(D_1^{\text{st}}) > z(D_2^{\text{st}}).$$

Now let D_1^d be a disjunction of $n D^{\text{st}}$ which have the same z as D_1^{st} . Then D_1^d is logically equivalent to a disjunction of $n \times z(D_1^{\text{st}}) D^{\text{ind}}$. Therefore, with (6-21) and (6-11):

$$(6-22) \quad S_B^{II}(D_1^d) > S_B^{II}(D_2^{\text{st}}).$$

Hence with (6-17)(b):

$$(6-23) \quad S_B^{II}(D_1^d) > S_{\text{th}}(g, t_2).$$

The $n D^{\text{st}}$ in D_1^d may be typical D^{st} for states of g similar to that in the period around t_1 ; only the wall between the parts A and B is in

different positions, while the volumes of A and B and the values of T_A , T_B , p_A , and p_B are the same in all cases. Then D_1^{st} says that g is in one of n states which have the same value of S_B , equal to $S_B(D_1^{\text{st}})$, hence $< S_B(D_2^{\text{st}})$, and the same value of S_{th} , equal to $S_{\text{th}}(g, t_1)$, hence $< S_{\text{th}}(g, t_2)$. Thus it follows from D_1^{st} that the S_{th} of g is smaller than $S_{\text{th}}(g, t_2)$; on the other hand, S_B^{II} is greater than for D_2^{st} (6-22), and greater than $S_{\text{th}}(g, t_2)$ (6-23). In this case, *the inequality between the values of S_B^{II} (6-22) is the inverse of that between the values of S_{th}* .

We found that S_B^{II} is not a purely physical concept. Now it is easy to see that it is not even a mixed concept with a physical component, but a purely logical concept. This is shown by the following example.

Example 3. Suppose we know that the gas g was in a state of thermodynamic equilibrium during a time interval around t_1 ; we know the volume of the vessel and the volumes of the parts separated by walls, the number N of molecules and the numbers for the parts, and the molecular mass m ; furthermore, a cell system Ω^u with V^u and K is specified. Now let us consider various descriptions; they are meant as typical descriptions of g for the interval around t_1 .

(3a) Suppose that D_1^{st} is given to us. Then we can calculate the value of $S_B(D_1^{\text{st}})$, say r_1 (a particular real number), and we can infer that S_{th} is likewise $= r_1$.

(3b) Suppose that instead of D_1^{st} we obtain only the statement " $S_B = r_1$." This means: "The (unspecified) D^{st} for g at t_1 is such that its S_B is r_1 ," because the (unextended) concept S_B applies only to D^{st} . From this statement we can still infer that $S_{\text{th}} = r_1$.

(3c) Suppose we are told: " $S_B^{\text{I}} = r_1$." S_B^{I} is applicable to any description giving a set of cell numbers N_i or a common value of S_B for several D^{st} , and the value of S_B^{I} for such a description is equal to the corresponding value of $S_B(D^{\text{st}})$. Therefore the present statement means the same as that in (3b).

(3d) Now we consider, in contrast, the situation with S_B^{II} . Let the statement " $S_B^{\text{II}} = r'_1$ " be given, with a specified real number r'_1 . S_B^{II} is defined for any disjunction D_m of m D^{ind} with any m ($1 \leq m \leq Z$). Therefore the given statement means: "There is a number m and an (unspecified) D_m which holds for g at t_1 and for which $S_B^{\text{II}} = r'_1$." Now we see from (6-11) that, since k is known and

V^μ , K , and N are supposed to be given, S_B^{II} depends merely on m ; and conversely from the given S_B^{II} -value r'_1 , we can determine the value of m , say m_1 [viz., $m_1 = e^{r'_1/k} \left(\frac{NK}{V^\mu} \right)^N$]. Suppose we find $m_1 = 1000$. Then the given statement means: "There is a disjunction of one thousand D^{ind} which holds for g at t_1 ." The statement does not tell us which D^{ind} belong to the disjunction; it says nothing about the content of these D^{ind} , the cell numbers or anything else. It says merely that the (unspecified) disjunction consists of one thousand D^{ind} ; this is a logical characteristic concerning logical strength or amount of information. The statement can also be formulated in this way: "The one (unspecified) D^{ind} which holds for g at t_1 belongs to an (unspecified) class of one thousand D^{ind} ." This holds obviously for every D^{ind} and for every possible state of g . It is a tautology, a purely logical truth. Therefore S_B^{II} is a *purely logical concept*. If it is to be introduced at all, the term "entropy" should not be used for it.

(3e) The situation is quite similar, if the statement " $S_B^{\text{IR}} = r'_1$ " is given, where S_B^{IR} is defined in terms of pr_B (see the earlier remark on (6-15)). From the given statement we can determine the probability of an (unspecified) description which holds for g at t_1 , say $pr_B = 0.0001$. Thus what we learn about g at t_1 is merely that its state can truly be described by an (unspecified) description which has the probability 0.0001. But this again holds for any possible state of the gas.

The strangeness of the results to which Method II leads becomes especially striking in the following case.

Example 4. Let D_1^S assign to g at t_1 a value of S_B with a certain precision, by specifying a small interval on the S_B -scale. We take again the D_1^{st} of Examples 1 and 2. D_1^{st} was assumed to be typical for g at t_1 . To make the example more concrete, let us assume that

$$(6-24) \quad S_B(D_1^{\text{st}}) = -10^7 k.$$

Let D_1^S say that the mean S_B for the period around t_1 is $-10^7 k \pm \delta$. Then D_1^S is equivalent to a disjunction of certain D^{st} , namely those whose S_B -values lie within the given interval, among them D_1^{st} . Since each D^{st} is a disjunction of certain D^{ind} , D_1^S is equivalent to a disjunction of a great number of D^{ind} , say m . Suppose that the size 2δ of the interval is such that $m = Z/10^6$; thus the interval covers one millionth of the number of all D^{ind} . Then $pr_B(D_1^S) = 10^{-6}$. So far we did not specify either V^μ or N ; let us assume for simplicity that

they happen to be equal. Then from (6-15), $S_B^{II}(D_1^S) = k \ln pr_B(D_1^S) = -13.8k$. Since S_B^{II} claims to be a concept of entropy and simply an extension of S_B , the result obtained may be formulated as follows:

(6-25) “The entropy (S_B^{II}) of the gas g at t_1 on the basis of the description D_1^S , which says that the entropy (S_B) of g at t_1 is near to $-10,000,000k$, is $-13.8k$.”

This result is, to say the least, rather strange. D_1^{st} was assumed to be a typical description of g for the specified period. Hence the mean value of S_B for g during the period is -10^7k . According to (6-17) (a), this is then also the value of S_{th} . D_1^S is likewise true; but the value of S_B^{II} for D_1^S is $-13.8k$. Obviously this value has no relation at all to thermodynamic entropy.

The application of Method I leads in this example to the only plausible result, which is here not a point value but an interval. D_1^S is a disjunction of certain D_i^{st} . For each of them, $S_B^I(D_1^{st}) = S_B(D_i^{st})$, and likewise, according to (6-8), for any partial disjunction consisting of D^{st} which have the same S_B -value as D_i^{st} . The D^{st} contained in D_1^S are those for which S_B is in the interval $-10^7k \pm \delta$; therefore S_B^I determines for D_1^S just this interval.

The analyses in this section have shown the following. If the extension of S_B is intended to maintain the character as a physical concept in statistical mechanics which is related to the thermodynamic concept of entropy, their Method II must be rejected and Method I must be applied.

7

Gibbs's Statistical Method

Summary. The state of the whole system of N elements, e.g., the gas g , is represented by a point U in the Mn -dimensional γ -space. A description D_1 of g at t_1 is represented by the point set R_1 in the γ -space, called the range of D_1 . Let the volume of R_1 be v_1 . Then the (probability) density $\rho_1(U)$ based on D_1 is $1/v_1$ at any point U in R_1 , and 0 outside of R_1 . Let D_2 be the description of g at a later time t_2 , derived from D_1 with the help of the laws of mechanics. The range of D_2 has the same volume as that of D_1 (Liouville's theorem), although a different location and shape.

The development of new method for statistical mechanics by Gibbs¹ constitutes an essential progress. We shall give a brief exposition of the essential features of his method, based chiefly on Tolman.² We shall, however, use some of our own terms and notations. This seems necessary because the customary formulations, both in words and in symbols, are almost always elliptical and therefore often not quite clear.

In the next section we shall explain Gibbs's definition of entropy and show that it is characterized by what we have called Method II. Then we shall indicate an alternative definition (§9) which uses instead Method I and thereby leads to a magnitude which is in accordance with the principles for purely physical concepts, (6-1) and (6-2).

In Boltzmann's method, each of the N molecules a_i of the gas g , characterized by the values of n magnitudes u_{i1}, \dots, u_{in} , is represented by a phase point in the n -dimensional μ -space. Gibbs represents the state of g itself, or its description D^{prec} , which is characterized by the Nn values $u_{11}, \dots, u_{1n}, u_{21}, \dots, u_{2n}, \dots, u_{N1}, \dots, u_{Nn}$, by a phase point in the Nn -dimensional phase space, called γ -space (gas-space). We shall use ' U ' as a variable for an ordered set of Nn values u_{11}, \dots, u_{Nn} as coordinates, and hence for a point in the γ -space. R^μ determines the total range R^γ in the γ -space; let V^γ be its (Nn -dimensional) volume ($V^\gamma = (V^\mu)^N$).

¹ J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (New Haven; Yale University Press, 1902).

² Richard C. Tolman, *The Principles of Statistical Mechanics* (London: Oxford University Press, 1938), Ch. vi.

Now the main new feature of Gibbs's method is as follows. If the precise description D_1^{prec} of the gas body g at the initial time t_1 were known, the precise description D_2^{prec} of g at a later time t_2 could be determined with the help of the laws of classical mechanics. In fact, however, actual observations (or hypothetical assumptions) supply only a weaker description of g at t_1 , say D_1 . This description determines, not a single point, but a region in the γ -space; this we call the *range* of D_1 , $R(D_1)$. Gibbs calls the class $R(D_1)$ of phase points, or the class of possible states of the gas g represented by these phase points, the representative *ensemble* of g at t_1 as described by D_1 .

An ensemble is characterized in Gibbs's method by a probability density function $\rho(U)$ over the γ -space. The probability on the basis of ρ that the phase point of the gas state represented by the ensemble belongs to a region R_i is:

$$(7-1) \quad pr(R_i) = \int_{R_i} \rho(U) dU.$$

('dU' is short for ' $du_{11} \cdots du_{Nn}$ '; the integral symbol stands for an Nn -fold integral.) Any density function ρ is normalized to 1:

$$(7-2) \quad \int \rho(U) dU = 1, \text{ where the integral covers the whole } \gamma\text{-space.}$$

Outside of the total range R^γ , ρ is taken to be 0. Therefore the integral (7-2) over R^γ is likewise 1. On the basis of any description D_i , the density function ρ_i based on D_i is taken to be 0 outside of $R(D_i)$. Therefore:

$$(7-3) \quad \int_{R(D_i)} \rho_i(U) dU = 1.$$

Let F be a physical magnitude $F(g, t)$ of g at time t , which is defined as a function of the Nn value u_{11}, \dots, u_{Nn} for g at t : $F(u_{11}, \dots, u_{Nn}; t)$. It may therefore be represented as a function over the γ -space: $F(U, t)$. If a density function $\rho(U, t)$ is given, the *ensemble mean* \bar{F} of F at t (in the sense of the probability mean, in other words, the expectation value) is defined as follows:

$$(7-4) \quad \bar{F}(g, t) = \text{Df} \int F(U, t) \rho(U, t) dU.$$

Gibbs takes this ensemble-mean $\bar{F}(g, t)$, based on the density function ρ for t determined by a description D_i , as the value we should reasonably expect for $F(g, t)$ if D_i represents all the data we have

(or assume). In the terminology of inductive logic, this means the estimate of $F(g, t)$ with respect to the observational evidence D_i :

$$(7-5) \quad \text{est}(F, g, t, D_i) = \bar{\bar{F}}(g, t).$$

The application of this method of estimation presupposes two postulates: (1) the laws of mechanics, (2) a probability assumption. The laws (1) are used only if the time t of the estimated value differs from the time t_i to which D_i refers; in this case the laws are needed for deriving $\rho(U, t)$ from $\rho(U, t_i)$ determined by D_i . The probability postulate is needed for the determination of $\rho(U, t)$ from D_i .

The following postulate is assumed, in accordance with Boltzmann's probability assumption (4-8):

(7-6) *Postulate of apriori density:*

- (a) Outside R^γ , $\rho = 0$.
- (b) Within R^γ , ρ is constant = ρ_0 .

Any description of g at t in our discussions is supposed to be such that it merely excludes some possible cases, i.e., some values of the n magnitudes ϕ_i , but does not say anything about the probabilities of the nonexcluded cases. Thus it may state, for example, intervals for some of the m magnitudes or for other magnitudes defined on their basis, but it does not give a probability distribution within any of these intervals. Therefore a description D_i is completely characterized by its range R_i . Then the postulate (7-6) yields:

(7-7) *Apriori probability pr_0 for a description D_i whose range R_i has the volume v_i :*

$$pr_0(D_i) = pr_0(R_i) = \int_{R_i} \rho_0(U) dU = v_i \rho_0.$$

Since for a tautology, whose range is R^γ with volume V^μ , pr_0 must be 1:

$$(7-8) \quad \rho_0 = \frac{1}{V^\mu}$$

Hence:

$$(7-9) \quad pr_0(D_i) = pr_0(R_i) = \frac{v_i}{V^\gamma}.$$

(7-10) Rule for the *aposteriori density* $\rho_i(U, t)$ based on a description D_i for t :

- (a) Outside R_i , $\rho_i(U, t) = 0$,
- (b) Within R_i , $\rho_i(U, t)$ is constant.

Hence with (7-3):

$$(7-11) \quad \text{For any point } U \text{ within } R_i, \rho_i(U, t) = \frac{1}{v_i}.$$

The *aposteriori or relative probability* of a description D_j with range R_j on the basis of D_i , where D_i and D_j refer to the same time t , is defined as follows:

$$(7-12) \quad pr(D_j | D_i) = \int_{R_j} \rho_i(U, t) dU.$$

This leads to the customary form:

$$(7-13) \quad pr(D_j | D_i) = \frac{pr_0(D_i \cdot D_j)}{pr_0(D_i)}.$$

Hence, if the range of the conjunction $D_i \cdot D_j$ (which is the intersection of R_i and R_j) has the volume v_{ij} :

$$(7-14) \quad pr(D_j | D_i) = \frac{v_{ij}}{v_i}.$$

Let D_1 be a description of g at time t_1 . Let the range of D_1 be R_1 with volume v_1 . Let D_2 be the description of g at a later time t_2 , as derived from D_1 according to the laws of mechanics. Let the range of D_2 be R_2 with volume v_2 . We assume for D_1 , but not for D_2 , that it is a description not involving probability. Hence we derive from D_1 the density function $\rho_1(U)$ according to (7-10) and (7-11). The laws of mechanics determine, for a given initial phase point U_1 at t_1 , the later phase point U_2 at t_2 . Therefore, from the initial density function $\rho_1(U)$, the density function $\rho_2(U)$ for t_2 can be derived. With respect to this situation, the following important results (7-15) and (7-16) hold, known as *Liouville's theorem* (1838):

$$(7-15) \quad v_1 = v_2.$$

This means that, although the range of the ensemble changes location and shape in the course of time, its volume remains unchanged.

(7-16) (a) Outside R_2 , $\rho_2(U)=0$.

(b) For every U within R_2 , $\rho_2(U)=\frac{1}{v_2}=\frac{1}{v_1}$
($=\rho_1$ within R_1).

Thus ρ_2 is constant in R_2 . Therefore D_2 is likewise a description not involving probability but completely characterized by its range.

8

Gibbs's Definition of Entropy for an Ensemble

Summary. On the basis of Boltzmann's cells in the μ -space, Gibbs defines cells Q^γ in the γ -space. If a "fine-grained" density $\rho_i(U)$ is given the "coarse-grained" density $P_i(Q_k^\gamma)$ is defined as the mean of ρ_i in the cell Q_k^γ . Supposedly as an analogue to Boltzmann's H , Gibbs's H_G^H is defined as the ensemble mean of $\ln P_i$. Gibbs's entropy concept S_G^H is then defined as $-kH_G^H$. Both concepts are applicable not only to D^a but to descriptions of any form. They use implicitly Method II (§6). For any D^{ind} or disjunction of several D^{ind} or any D^a or disjunction of D^a , S_G^H has, aside from an additive constant the same value as S_B^H . It is shown that S_G^H is in general, except for D^a , not in agreement with thermodynamic entropy S_{th} . Gibbs states a theorem which says that, under certain conditions, S_B^H increases in the course of time. It is shown that this increase has nothing to do with the increase of S_{th} and therefore Gibbs's theorem is not a statistical counterpart to the second law of thermodynamics, as it is supposed to be.

The aim of Gibbs's definition of entropy was essentially the same as that of Boltzmann's, viz., to provide a statistical concept corresponding to the thermodynamic concept of entropy S_{th} . But Gibbs's concept is based on his new statistical method and is therefore defined with respect to ensembles in the γ -space.

Gibbs uses for his definition a cell system Ω^γ in the γ -space, based upon Boltzmann's cell system Ω^a in the μ -space. Ω^a was a division of the total range R^a into K μ -cells, each having the μ -volume $v^a = \frac{V^a}{K}$. This determines the system Ω^γ , which divides

$$R^\gamma \text{ into } Z = K^N \text{ } \gamma\text{-cells } Q^\gamma \text{ of equal } \gamma\text{-volume } v^\gamma = (v^a)^N = \frac{(V^a)^N}{K^N} =$$

$\frac{V}{Z}$. A D_i^{ind} determines for each of the N molecules its μ -cell; therefore its range in the γ -space is one γ -cell Q_i^γ . Let D_i^{st} have the μ -cell numbers N_1, \dots, N_k . Then the number of the D^{ind} corresponding to D_i^{st} is $z(D_i^{st}) = \frac{N!}{N_1! \dots N_k!}$. The range of D_i^{st} in the γ -space consists of the z γ -cells representing those D^{ind} .

If any density function $\rho_i(U)$ is given, e.g., determined by a description D_i , then, on the basis of this "fine-grained" density $\rho_i(U)$, the corresponding "coarse-grained" density $P_i(U)$ at any

point U is defined as the mean of ρ_i over the γ -cell to which U belongs:

For any point U_j within a γ -cell Q_k^γ ,

$$(8-1) \quad P_i(U_j) =_{\text{Df}} \frac{1}{v^\gamma} \int_{Q_k^\gamma} \rho_i(U) dU$$

Note that P_i , in distinction to ρ_i , is dependent on the cell system Ω^γ . Since P_i has the same value for all points in any given γ -cell, we may regard it as a function of the cells: $P_i(Q_k^\gamma)$.

Let D_j^{ind} have the range Q_j^γ ; then we see from (7-12):

$$(8-2) \quad \text{pr}(D_j^{\text{ind}} | D_i) = P_i(Q_j^\gamma) v^\gamma.$$

Corresponding to Boltzmann's H -function, Gibbs defines a certain function as an ensemble mean; we denote it by H_G^{II} (because it implicitly uses Method II as we shall see):

$$(8-3) \quad H_G^{\text{II}}(D_i) =_{\text{Df}} \sum_n [P_i(Q_n^\gamma) \ln P_i(Q_n^\gamma)] v^\gamma,$$

where the sum runs through all γ -cells. Hence with (8-1), since $\ln P_i(U)$ has the same value for all points in a cell:

$$(8-4) \quad H_G^{\text{II}}(D_i) = \int_{R^\gamma} \rho_i(U) \ln P_i(U) dU.$$

Thus H_G^{II} is the ensemble mean of $\ln P_i$ (see (7-4)), not that of H ! [Therefore the notation \bar{H} sometimes used for H_G^{II} is quite misleading; it obviously violates the convention (7-4) concerning the symbol of the mean.]

Gibbs's entropy concept, which we shall denote by ' S_G^{II} ', is then defined in analogy to (4-6):

$$(8-5) \quad S_G^{\text{II}}(D_i) =_{\text{Df}} -kH_G^{\text{II}}(D_i).$$

[Actually Gibbs used the term 'temperature' not for T but for $\theta = kT$, and consequently the term 'entropy' not for S_{th} but for S_{th}/k and therefore took as its statistical counterpart not S_G^{II} but S_G^{II}/k . Since the difference is inessential, we explain Gibbs's method by translating it into the more customary terminology.]

We shall now examine Gibbs's concepts H_G^{II} and S_G^{II} . They are, like Boltzmann's concepts H and S_B , dependent upon the choice of a cell system, but, in distinction to Boltzmann's original concepts, they have the advantage that they are not restricted to D^{st} but applicable to any form of description with any region in the γ -space as its range.

Let D_i be a disjunction of $m D^{\text{ind}}$ (not necessarily corresponding to the same D^{st}). Let these D^{ind} be represented by the m γ -cells Q_1, \dots, Q_m . Then the range $R(D_i)$ consists of these m γ -cells and thus has the volume mv^γ . Therefore the density ρ_i based on D_i has for every point of this range the value $1/(mv^\gamma)$ (see (7-11)). Thus this is likewise the value of P_i for each of the m γ -cells. Hence from (8-3):

$$(8-6) \quad H_G^{\text{II}}(D_i) = -\ln(mv^\gamma).$$

$$(8-7) \quad S_G^{\text{II}}(D_i) = k(\ln m + \ln v^\gamma).$$

Hence with (6-11) (b):

$$(8-8) \quad S_G^{\text{II}}(D_i) = S_B^{\text{II}}(D_i) + kN \ln N.$$

Thus $S_G^{\text{II}}(D_i)$ is, aside from an additive constant, the same as $S_B^{\text{II}}(D_i)$. Therefore S_G^{II} possesses likewise the essential features of Method II.

We shall now examine some special cases of descriptions of the specified form D_i .

Case 1. Let D_i be a D_i^{st} ; then $m = z(D_i^{\text{st}})$.

Hence:

$$(8-9) \quad H_G^{\text{II}}(D_i^{\text{st}}) = -(\ln z(D_i^{\text{st}}) + \ln v^\gamma).$$

$$(8-10) \quad S_G^{\text{II}}(D_i^{\text{st}}) = k(\ln z(D_i^{\text{st}}) + \ln v^\gamma).$$

For a fixed, sufficiently large N and a fixed cell system, these values differ from those of H and S_B , respectively, only by an additive constant (see (4-5) and (4-7)). Thus in the case of a D^{st} , Gibbs's concept of entropy is like Boltzmann's, in agreement with S_{th} under the conditions previously mentioned (§5). However, we shall see that this agreement holds only for D^{st} .

Case 2. Let D_i be a D^{ind} ; then $m = 1$ and hence:

$$(8-11) \quad H_G^{\text{II}}(D_i^{\text{ind}}) = -\ln v^\gamma.$$

$$(8-12) \quad S_G^{\text{II}}(D_i^{\text{ind}}) = k \ln v^\gamma.$$

Thus S_G^{II} has the same value for all D^{ind} , irrespective of the physical properties of the state of the gas as described in the D^{ind} .

As a consequence of the use of Method II, S_G^{II} like S_B^{II} has in this case no relation to S_{th} ; moreover, it is not a physical concept but a purely logical one.

Case 3. Let D_i be a disjunction D_1 of $p_1 D^{\text{st}}$ such that in any two of these D^{st} the cell numbers N_i for any μ -cell Q_i are either equal or nearly equal. Then all these D^{st} have practically the same value of z , say z_1 , and therefore also of S_B , say $S_{B,1}$, and of S_G^{II} (8-10), say $S_{G,1}^{\text{II}} = k(\ln z_1 + \ln v')$. Suppose that the conditions for agreement between S_B and S_{th} stated in §5 are satisfied; the $p_1 D^{\text{st}}$ may, for example, be those which are compatible as typical D^{st} for an interval around t_1 with some measurements of temperature and pressure made at various places in the gas. More specifically, let us take the model explained at the end of §5; the vessel is divided by walls perforated by small holes. At any time during the process, the gas g is in each part of the vessel in a state of near-equilibrium. We have $m = m_1 = p_1 z_1$, and hence with (8-7): $S_G^{\text{II}}(D_1) = k[\ln(p_1 z_1) + \ln v'] = S_{G,1}^{\text{II}} + k \ln p_1$. Suppose that an interval around t_2 is only a short time later than the first and still within the period of equalization. Suppose that in this interval similar measurements are made, with slightly different results but much greater precision. Let D_2 be again the disjunction of the corresponding D^{st} ; let their number be p_2 . Let z_2 , $S_{B,2}$, and $S_{G,2}^{\text{II}}$ be the values which hold (practically) for each of these D^{st} . Then: $S_G^{\text{II}}(D_2) = k[\ln(p_2 z_2) + \ln v']$. Because of the equalization, we have $S_{\text{th}}(g, t_2) > S_{\text{th}}(g, t_1)$; therefore $S_{B,2} > S_{B,1}$, $S_{G,2}^{\text{II}} > S_{G,1}^{\text{II}}$, and $z_2 > z_1$. We assume that the holes are very small and therefore the rate of equalization very low and that $t_2 - t_1$ is small. Therefore the increase in z will be small. On the other hand, we assume that the precision of the measurements at t_2 is much greater than that at t_1 and therefore p_2 is much smaller than p_1 , so much smaller that $m_2 = p_2 z_2 < m_1 = p_1 z_1$. Hence:

$$(8-13) \quad S_G^{\text{II}}(D_2) < S_G^{\text{II}}(D_1).$$

Thus for the descriptions D_1 and D_2 of the kind described, which are (practically) typical descriptions of g for t_1 and t_2 respectively, we find a decrease in S_G^{II} in spite of an increase in S_{th} . This shows that in cases of this kind Gibbs's concept S_G^{II} is not in accord with S_{th} .

Case 4. The application of S_G^{II} to the description D_1^s in Example 4 in §6 would lead to a result which is analogous to (6-25) and just as strange.

Thus we are led to the conclusion that Gibbs did not reach his aim of constructing a statistical concept corresponding to S_{th} . The correspondence holds for his concept only in the special case of D^s . And in Gibbs's method the concept cannot be restricted to D^s ; this will be seen in the discussion of his H -theorem.

Gibbs has stated for his function H_G^H a theorem which is regarded as an analogue to Boltzmann's H-theorem (5-2):

- (8-14) *Gibbs's "generalized H-theorem"* for his function H_G^H . Let D_1 be set up as a typical description for g on the basis of measurements made during a time interval around t_1 , with given N and the total energy lying in a small interval around E . Then H_G^H will diminish and hence S_G^H increase in the course of time until the representative ensemble is nearly evenly spread out over the γ -cells of the given E -interval.

Various arguments have been proposed to prove this theorem or to make it at least plausible on the basis of various additional assumptions (see, e.g., Tolman, op. cit., pp. 169-179).

We shall give here a simple proof of a partial assertion of (8-14) which gives a comparison between the states at t_1 and some later time t_2 . Our main purpose is to clarify the physical meaning of the decrease in H_G^H .

Let D_1 be a disjunction of $m_1 D^{ind}$ which have the same or nearly the same μ -cell numbers N_i , typical for the gas g during an interval around t_1 . If D_1 is based on measurements, it may, for example, be a D^s , as in Case 1 discussed above (with $m_1 = z_1$), or a disjunction of $p_1 D^s$ as in Case 3 (with $m_1 = p_1 z_1$); however, this is not necessary for the following proof. Let D_2 be that description of g at t_2 which is derived from D_1 under the assumption of constant energy (see the explanation at the end of §7). Let R_1 be the range of D_1 in the γ -space, and R_2 that of D_2 . R_1 consists of m_1 γ -cells; therefore its volume is $v_1 = m_1 v^\gamma$. The density ρ_1 for D_1 is $1/(m_1 v^\gamma)$ within R_1 and 0 outside. According to Liouville's theorem ((7-15) and (7-16)), the volume v_2 of R_2 is likewise $m_1 v^\gamma$, and the density ρ_2 for D_2 is $1/(m_1 v^\gamma)$ within R_2 and 0 outside. Suppose that R_2 extends into m_2 cells, $Q_{2,n}^\gamma$ ($n = 1, \dots, m_2$), and that the volume of the part of $Q_{2,n}^\gamma$ covered by R_2 is $q_n v^\gamma$. Hence for every $n (= 1, \dots, m_2)$, $0 < q_n \leq 1$. Since $v_2 = m_1 v^\gamma$:

$$(8-15) \quad \sum_{n=1}^{m_2} q_n = m_1.$$

Since the ensemble changes its shape all the time, it seems quite plausible to assume that at t_2 , although its volume is still $m_1 v^\gamma$, it does not exactly consist of entire cells only; in other words, that at least one of the cells $Q_{2,n}^\gamma$ is not completely covered by it:

(8-16) Assumption.

- (a) For at least one n , $0 < q_n < 1$; hence
- (b) $m_2 > m_1$.

Let P_2 be the coarse-grained density for D_2 ; then from (8-1):

$$(8-17) \quad \text{For } n = 1, \dots, m_2, \quad P_2(Q_{2,n}^\gamma) = \frac{1}{v^\gamma} \frac{q_n v^\gamma}{m_1 v^\gamma} = \frac{q_n}{m_1 v^\gamma}; \quad \text{for all other cells, } P_2 = 0.$$

Therefore with (8-3) and (8-15):

$$(8-18) \quad H_G^H(D_2) = \sum_{n=1}^{m_2} \left[\frac{q_n}{m_1 v^\gamma} n \frac{q_n}{m_1 v^\gamma} \right] v^\gamma \\ = \frac{1}{m_1} \sum [q_n \ln q_n] - \ln (m_1 v^\gamma).$$

From (8-16):

$$(8-19) \quad H_G^H(D_1) = -\ln (m_1 v^\gamma).$$

$$(8-20) \quad \begin{aligned} (a) \quad H_G^H(D_1) - H_G^H(D_2) &= -\frac{1}{m_1} \sum_{n=1}^{m_2} [q_n \ln q_n], \\ (b) &\quad > 0. \quad (\text{From (8-16) (a).}) \end{aligned}$$

Thus H_G^H decreases from t_1 to t_2 .

Gibbs's theorem is supposed to be, like Boltzmann's H -theorem, a statistical counterpart to the second law of thermodynamics. However, there are striking differences between Gibbs's and Boltzmann's theorems in two points. First, at least the partial assertion (8-20) (b) does not, like Boltzmann's theorem, involve probability; it states that H_G^H decreases and S_G^H increases from t_1 to t_2 with certainty, once the assumption (8-16) (a) is granted. And we see from (8-20) (a) that the amount by which H_G^H decreases depends merely upon the number m_2 of cells and the fractions q_n of these cells covered by R_2 ; the amount is independent of the physical properties of these cells and, in particular, of the values of H and of S_{th} for the gas states represented by these cells.

Thus the decrease in H_G^H has nothing whatever to do with a decrease in H and an increase in S_{th} . Therefore *Gibbs's theorem is not a generalization of Boltzmann's H-theorem, and has nothing to do with the second law of thermodynamics*. This will become still clearer by the consideration of two special cases.

Case 5. Suppose that g is in a state of local near-equilibrium both at t_1 and at t_2 . Let $S_B(t_1)$ be the mean S_B for an interval around t_1 and $S_B(t_2)$ the mean S_B for a very short interval around t_2 . Suppose that $S_B(t_2)$ is smaller than $S_B(t_1)$ by a very small amount, but that this amount and the second interval are still large enough for measurement. According to the statistical conception, this case is not impossible although its probability is extremely small, so that for all practical purposes we may disregard the case. We consider now this case, not for practical purposes, but only in order to clarify the meaning of S_G^H . In this exceptional case, the mean S_B and therefore the measurable S_{th} decrease, but according to (8-20) (b), S_G^H nevertheless increases. This shows again that S_G^H has nothing to do with S_{th} . In this case most of the new cells into which the ensemble has spread by t_2 represent states with lower values of S_B and of S_{th} . S_G^H in no way reflects this fact; it merely reflects the fact of spreading.

Case 6. Let D_1 be a typical D^{st} with the minimum value of H for the given E , and therefore the maximum of S_B and of S_{th} . Thus g is at t_1 in a state of (total, not only local) statistical and thermodynamic equilibrium. While Boltzmann's theorem contains the condition that $H(D_1^{st})$ be above the mean and thus does not apply to the present case, Gibbs's theorem is not restricted in this respect. Obviously, in this case, S_B cannot possibly increase and will most probably even decrease (by a very small amount); and S_{th} cannot increase. Even Gibbs's own function S_G^H has in this case for D_1 the maximum value among all D^{st} for the given E . (See the remark on (8-10) above). Nevertheless, according to (8-20) (b), S_G^H is higher for D_2 than for D_1 (this is possible only because D_2 is not a D^{st}). It may even be considerably higher, and thus the discrepancy between S_G^H and S_{th} may be considerable. [This is seen as follows. H_G^H has its minimum in the given E -interval if m_2 is the number m_E of cells for the E -interval and all q_n are equal and hence, with (8-15), $=m_1/m_E$. If $H_G^H(D_2)$ has this minimum value, $H_G^H(D_1) - H_G^H(D_2)$ is, according to (8-20) (a), $=\ln(m_E/m_1)$. Although those D^{st} whose S_B is equal or near the maximum S_B for the E -interval have together a high probability and cover a majority of the m_E cells, nevertheless, since

the number of these D^* is large, our D_1 alone, which is one of them, covers only a small part of the cells. Therefore m_1/m_E is a small fraction.]

A short remark only about the development of the ensemble after the time t_2 : Suppose that at a later time t_3 the ensemble has spread into m_3 cells with q -values $q_{n'}$ ($n' = 1, \dots, m_3$). Then a sufficient and necessary condition for H_G^{II} being smaller at t_3 than at

t_2 is that $\sum_{n=1}^{m_2} \left(q_n \ln \frac{1}{q_n} \right) < \sum_{n=1}^{m_3} \left(q_{n'} \ln \frac{1}{q_{n'}} \right)$; hence that there is an increase from t_2 to t_3 in the degree of disorder of the distribution of the ensemble points over the γ -cells. This condition means roughly a spreading into still more cells and a more even distribution of the q -values. It does *not* mean an increase in the degree of disorder of the distribution of the phase points of the molecules among the μ -cells, which would be equivalent to an increase in S_B . Thus the logical situation with Gibbs's theorem is as follows. By an analysis of the laws of mechanics (similar to the analysis which led to Liouville's theorem) the physicists regard it as plausible that the following assumption holds with very great probability (i.e., in the great majority of all possible cases)!

(A) If at t_1 the degree of disorder of the distribution of a given ensemble among the γ -cells is not near to its maximum (which is the state in which the ensemble covers an equal fraction q of every γ -cell corresponding to the given E), then at a later time t_2 the degree of disorder will be higher than at t_1 .

From (A) the following can be inferred:

(B) If at t_1 H_G^{II} for a given ensemble is not near to its minimum value for the given E , then at a later time t_2 H_G^{II} will be lower. (B) is essentially Gibbs's theorem. (B) is mathematically deducible from (A). Now our discussions have shown that from the mere statement of the decrease of H_G^{II} in (B) nothing can be inferred about a change in S_B or in S_{ch} .

9

An Alternative Definition of Entropy for an Ensemble.

Summary. We define H_G^I as the ensemble mean of H^I and S_G^I as $-kH_G^I$. Then S_G^I is the ensemble mean of our S_B^I (§6). We propose S_G^I as an alternative to Gibbs's S_G^{II} . In contrast to the latter, S_G^I is in agreement with S_{th} not only for D^st but also for more specific or less specific descriptions. A generalized H -theorem (9-12) involving our concepts H_G^I and S_G^I is formulated, which follows from the same assumptions as Boltzmann's H -theorem. This theorem is, in contrast to Gibbs's theorem (8-14), a statistical counterpart to the second law of thermodynamics.

We shall now construct an alternative concept which, in distinction to Gibbs's S_G^{II} , fulfills the following conditions: (1) it satisfies the principles (6-1) and (6-2) and is therefore a purely physical concept; (2) it is in agreement with S_{th} or with the mean of S_{th} in an ensemble (under the same conditions as S_B , §5); (3) it leads to a theorem which is an analogue to Boltzmann's H -theorem and a statistical counterpart to the second law. The property (1) is assured by the use of Method I (§6); we therefore denote the concept by ' S_G^I '. Although it is not Gibbs's concept, we retain the subscript ' G ', because we follow Gibbs in applying the concept to a representative ensemble in the γ -space.

In §6 we defined H^I for a D_i^{prec} as an extension of Boltzmann's H in accordance with Method I, with respect to a given cell-system (6-9). We assign now the same value of H^I to the phase point U_i which represents D_i^{prec} in the γ -space:

(9-1) $H^I(U_i) =_{Df} H(D_i^{st})$, where D_i^{st} is that D^st to whose range in the γ -space the point U_i belongs.

Hence from (4-4) (a):

(9-2) $H^I(U_i) = \sum (N_j \ln N_j) - \ln v^\gamma$, where the N_j ($j = 1, \dots, K$) are the μ -cell numbers of D_i^{st} .

Now we define H_G^I for any description D_i as the ensemble mean of $H^I(U)$, in accordance with (7-4):

(9-3) $H_G^I(D_i) =_{Df} \int H^I(U) \rho_i(U) dU$, where ρ_i is the density function determined by D_i .

Since the aim is to construct an analogue to Boltzmann's H -function, it seems more natural to take the mean of H^I than that of $\ln P_i$, as Gibbs does. [The notation \bar{H} , which is sometimes used for Gibbs's H_G^{II} , could correctly be applied to our concept; or, more specifically, the notation ' \bar{H}^I .'] For all points U of a γ -cell Q_n^γ , $H^I(U)$ has the same value $H^I(Q_n^\gamma)$; therefore:

$$(9-4) \quad H_G^I(D_i) = \sum_n [H^I(Q_n^\gamma) P_i(Q_n^\gamma)] v^\gamma,$$

where Σ extends over all γ -cells.

Now we define our entropy concept in analogy to (8-5):

$$(9-5) \quad S_G^I(D_i) =_{\text{df}} -k H_G^I(D_i).$$

Thus S_G^I is the ensemble mean of S_B^I , which is our extension (6-4) of Boltzmann's S_B .

Let D_i be a disjunction of m descriptions D_n^{ind} ($n = 1, \dots, m$). We found earlier (§8) that, for each of the m γ -cells Q_n^γ in the range of D_i , $P_i = 1/(mv^\gamma)$ and otherwise $P_i = 0$. Hence with (9-4):

$$(9-6) \quad H_G^I(D_i) = \frac{1}{m} \sum_{n=1}^m H^I(Q_n^\gamma).$$

$$(9-7) \quad S_G^I(D_i) = -\frac{k}{m} \sum_{n=1}^m H^I(Q_n^\gamma) = \frac{1}{m} \sum_{n=1}^m S_B^I(D_n^{\text{ind}}).$$

Thus $S_G^I(D_i)$ is the mean of S_B^I for the disjunctive components in D_i .

In order to compare our concepts with those of Gibbs, let us examine the results for the Cases 1 to 4 in §8.

Case 1. Let D_i be a D_i^{st} . Then from (6-5) and (6-6):

$$(9-8) \quad H_G^I(D_i^{\text{st}}) = H(D_i^{\text{st}}),$$

and

$$(9-9) \quad S_G^I(D_i^{\text{st}}) = S_B(D_i^{\text{st}}).$$

Thus in the case of a D_i^{st} , our S_G^I coincides exactly with Boltzmann's S_B . We saw that S_G^{II} does essentially the same (approximately, and with another additive constant; these differences may be regarded as inessential).

Case 2. Let D_i be a D^{ind} ; hence $m = 1$. Let D_i^{ind} be represented by the cell $Q_i^{\text{?}}$. Then:

$$(9-10) \quad H_G^I(D_i^{\text{ind}}) = H^I(Q_i^{\text{?}}).$$

$$(9-11) \quad S_G^I(D_i^{\text{ind}}) = S_B^I(D_i^{\text{ind}}) = S_B(D_i^{\text{st}}), \text{ where } D_i^{\text{st}} \text{ is the } D^{\text{st}} \text{ corresponding to } D_i^{\text{ind}}.$$

Thus for a D^{ind} , S_G^I coincides with our extension S_B^I of S_B , and with the value of S_B for the corresponding D^{st} . Since Boltzmann's $S_B(D^{\text{st}})$ and our $S_G^I(D^{\text{ind}})$ are in agreement with S_{th} (under the conditions stated in §5), the same holds for $S_G^I(D_i^{\text{st}})$ and $S_G^I(D_i^{\text{ind}})$. We found that S_G^{II} in this case has no relation to S_{th} .

Case 3. Let D_i be a disjunction D_1 of $p_1 D^{\text{st}}$ with nearly equal N_i . These D^{st} are assumed to be compatible with measurements made at t_1 . Similarly, let D_2 be a disjunction of $p_2 D^{\text{st}}$, related to measurements of much greater precision at t_2 . It was assumed that p_2 was much smaller than p_1 so that even $p_2 z_2 < p_1 z_1$, although $z_2 > z_1$. However, in distinction to S_G^{II} , the value of $S_G^I(D_1)$ is independent of p_1 ; it is $= S_{G,1}^I$, i.e., $S_G^I(D_1^{\text{st}})$, where D_1^{st} is any of the D^{st} in D_1 ; hence $= S_{B,1}^I$, i.e. $S_B(D_1^{\text{st}})$; hence $= S_{\text{th}}(g, t_1)$. Similarly, $S_G^I(D_2) = S_B(D_2^{\text{st}}) = S_{\text{th}}(g, t_2)$. Thus S_G^I is at both times in agreement with S_{th} and increases from t_1 to t_2 just like S_{th} . In contrast, we found that, due to the influence of the numbers p_1 and p_2 (which are characteristic of the precision of the measurements at t_1 and t_2 , respectively, but have nothing to do with the values obtained by the measurements) $S_G^{II}(D_2) < S_G^{II}(D_1)$ (8-13); thus the change in S_G^{II} is in the opposite direction to that in S_{th} .

Case 4. Let D_i be the description D_1^S in Example 4 in §6. D_1^S says that the mean S_B for the given period of time is near to -10^7 k . We see from (9-7) that $S_G^I(D_1^S) = -10^7 \text{ k}$, which is, just like S_B , in agreement with S_{th} . In contrast, Gibbs' concept leads to the paradoxical result $S_G^{II}(D_1^S) = -13.8 \text{ k}$, a glaring divergence from S_{th} .

Thus the result of the comparison between S_G^I and S_G^{II} is as follows. In the case of a D^{st} both concepts are in agreement with S_{th} (under the conditions of equilibrium and typical description). However, for any description which is either more specific than a D^{st} , e.g., a D^{ind} , or less specific, e.g., a disjunction of several D^{st} with equal or nearly equal S_B , S_G^I is still in agreement with S_{th} , while S_G^{II} deviates from S_{th} , sometimes by a very large amount.

We shall now formulate an analogue to Boltzmann's H -theorem (5-2) for our function H_G^I and consider the question of its validity. Let D_1^{st} be a typical description of an isolated gas body g at the time t_1 with total energy E ; let D_2 and D_3 be derived from D_1^{st} according to the laws of mechanics (§7) for the later time points t_2 and t_3 , respectively ($t_1 < t_2 < t_3$). Let \bar{H}_E be the mean value of

$H(D^{st})$ for the given E . [$\bar{H}_E = \sum_{i=1}^m [H(D_i^{st}) z(D_i^{st})] / \sum_{i=1}^m z(D_i^{st})$. In virtue of (4-9), where m is the number of D^{st} for the given E .]

(9-12) *Generalized H-theorem* for H_G^I . The following relations hold with overwhelming probability:

- (a) If $H(D_1^{st})$ is considerably higher than \bar{H}_E , $H_G^I(D_1^{st}) > H_G^I(D_2)$ and therefore $S_G^I(D_1^{st}) < S_G^I(D_2)$.
- (b) If $H_G^I(D_2)$ is considerably higher than \bar{H}_E , $H_G^I(D_2) > H_G^I(D_3)$ and therefore $S_G^I(D_2) < S_G^I(D_3)$.

We shall now show, that, if the assumptions underlying Boltzmann's theorem are valid (we do not decide whether they are or not), then (9-12) is valid. Let R_1 be the range of D_1^{st} in the γ -space, R_2 that of D_2 , and R_3 that of D_3 . Let $D_{1,i}^{prec}$ (with variable i) be any one of those D^{prec} which correspond to D_1^{st} . Let $D_{2,i}^{prec}$ and $D_{3,i}^{prec}$ be derived from $D_{1,i}^{prec}$ for t_2 and t_3 , respectively. Let $D_{2,i}^{st}$ and $D_{3,i}^{st}$ be the S^{st} corresponding to $D_{2,i}^{prec}$ and $D_{3,i}^{prec}$, respectively. Let $U_{1,i}$, $U_{2,i}$, and $U_{3,i}$ be the phase points of $D_{1,i}^{prec}$, $D_{2,i}^{prec}$ and $D_{3,i}^{prec}$, respectively. ($U_{1,i}$ runs through R_1 , $U_{2,i}$ through R_2 , and $U_{3,i}$ through R_3 .) From (9-1):

$$(i) \quad H^I(U_{1,i}) = H(D_1^{st}), \quad H^I(U_{2,i}) = H(D_{2,i}^{st}), \quad \text{and} \\ H^I(U_{3,i}) = H(D_{3,i}^{st}).$$

Boltzmann's theorem says in effect that the following relations hold with great probability if the conditions about \bar{H}_E in (9-12) (a) and (b) are fulfilled:

$$(ii) \quad H(D_{1,i}^{st}) > H(D_{2,i}^{st}) \quad \text{and} \quad H(D_{2,i}^{st}) > H(D_{3,i}^{st});$$

hence with (i):

$$(iii) \quad H^I(U_{1,i}) > H^I(U_{2,i}) \quad \text{and} \quad H^I(U_{2,i}) > H^I(U_{3,i}).$$

As earlier stated (§5), the reasoning leading to Boltzmann's theorem is not a rigorous proof; it rather shows that, if certain highly plausible assumptions are made, the relation asserted in this theorem holds in the average over a great number of possible cases. Thus the result of the reasoning can be formulated, more appropriately than by (ii), by saying that (iii) holds in the average for the ensemble in question. [This is explained, e.g., by Tolman, op. cit., pp. 146 f.] Now H_G^I was defined in (9-3) as the ensemble mean of $H^I(U)$. Therefore:

- (iv) $H_G^I(D_1^{st})$ is the ensemble mean of $H^I(U_{1,i})$,
- (v) $H_G^I(D_2)$ is the ensemble mean of $H^I(U_{2,i})$,
- (vi) $H_G^I(D_3)$ is the ensemble mean of $H^I(U_{3,i})$.

Thus the more appropriate formulation is as follows:

- (vii) With great probability, (a) $H_G^I(D_1^{st}) > H_G^I(D_2)$, and
- (vii) (b) $H_G^I(D_2) > H_G^I(D_3)$.

This is the theorem (9-12).

[Although we do not analyze in detail the reasoning leading to Boltzmann's theorem, some remarks may be made to indicate the weakest point in this reasoning. (Tolman, op. cit., pp. 148-152, discusses essentially the same point in greater detail and emphasizes the lack of proof for the assumptions involved. Unfortunately, his explanations and formulas there and at many other places throughout his chapter on H are not quite clear and are sometimes even ambiguous because he speaks of H "for a system s at time t " without explicitly specifying the description of s at t for which H is meant and without indicating how H is defined for other than statistical descriptions.) This point concerns part (b) of (9-12). At first glance, (b) may appear as perfectly analogous to and deducible from (a). But this is not the case, because D_1^{st} is a D^{st} , while D_2 is not. Let D'_2 be the disjunction of all D^{st} whose ranges overlap with R_2 , among them $D_{2,i}^{st}$. Hence the range R'_2 of D'_2 includes R_2 . Let D'_3 be derived from D'_2 for t_3 . Hence the range R'_3 of D'_3 includes R_3 . Let $(D_{2,i}^{st})_3$ be derived from $D_{2,i}^{st}$ for t_3 . Then (a), although not applicable to D_2 , can be applied to each D^{st} in D'_2 , e.g., $D_{2,i}^{st}$. Let us assume that (9-12) (a) is valid. Then:

- (viii) If $H(D_{2,i}^{st})$ is considerably higher than \bar{H}_E ,

$$H_G^I(D_{2,i}^{st}) > H_G^I(D_{2,i}^{st})_3,$$

and analogously for every other D^{st} in D'_2 . Therefore, an analogous relation holds for the means:

- (ix) If $H_G^I(D'_2)$ is considerably higher than \bar{H}_E ,

$$H_G^I(D'_2) > H_G^I(D'_3).$$

This is not yet (b). It refers to D'_2 and D'_3 , while (b) refers to D_2 and D_3 . But we can infer (b) from (ix) and hence from (a) if we make the following assumption:

- (x) With great probability, H_G^I has approximately equal values for D_2 and D'_2 , and likewise for D_3 and D'_3 .

This assumption can be made plausible by considering that H_G^I is the ensemble mean of $H^I(U)$ and the latter is constant within the range of any D^{st} and changes only very little from any such range to an adjacent one. However, the plausibility of (x) is weaker than that of the assumptions needed for part (a).

The assumptions needed for theorem (9-12) are the same as those needed for Boltzmann's theorem. Their necessity for the latter was discovered by Gibbs and his followers when they tried to reformulate Boltzmann's theorem and the reasoning leading to it in Gibbs's language. The refinement consisted chiefly in replacing Boltzmann's references to the values of certain magnitudes for a single physical system by references to the ensemble means of these magnitudes. Our theorem (9-12) has thus the same status as Boltzmann's. It might indeed be regarded as merely a practically equivalent formulation of the latter in Gibbs's ensemble language.

Let us assume that, at any moment during the whole time considered, g fulfills the condition of local near-equilibrium stated earlier (§5). Then S_B^I is in accord with S_{th} . S_G^I is the ensemble mean of S_B^I and thus is, in contrast to S_G^{II} , in accord with the average of S_{th} for the ensemble. Therefore, theorem (9-12) is, like Boltzmann's, but in contrast to Gibbs's theorem (8-14), a *statistical counterpart to the second law of thermodynamics*.

10

Entropy and Amount of Information

Summary. Let \inf be the amount of information (either in the statistical sense (10-1) or in the semantical sense (10-4), which have the same value under certain conditions). For any D^* , $S^* = -\inf$ (10-12); a similar relation holds between S_B and \inf (10-13). It is sometimes said that this relation holds generally, for all kinds of descriptions. However, this is the case only for those extended concepts which are formed by Method II, viz., S_B^{II} (6-11) and S^{*II} (10-15). We saw earlier (§6) that a concept of this kind is not a physical but a logical concept; now we see just which logical property it measures, viz. the negative amount of information. The general statement does not hold for the concepts formed by Method I. Szilard solved the paradox of Maxwell's demon by pointing out that the demon, in order to measure the velocity of an approaching molecule, must dissipate some energy and thereby increase the entropy. This is correct, but the conclusion drawn by later authors that in this process negative entropy is transformed into information and vice versa seems rather questionable.

In the mathematical theory of communication,¹ it is customary to define the amount of information in the following way. A finite set of messages or statements D_i is given, with probabilities $\text{pr}(D_i)$ assigned to them. We shall use here the term 'statistical information' and the symbol ' \inf_{stat} ' in order to distinguish this concept from another one to be introduced soon. The definition is as follows ('Log' denotes the logarithm to base 2):

$$(10-1) \quad \inf_{\text{stat}}(D_i) =_{\text{def}} -\text{Log} \text{pr}(D_i) = \text{Log} \frac{1}{\text{pr}(D_i)}.$$

In order to apply this concept, we make use of Boltzmann's probability assumption (4-8):

$$(10-2) \quad \text{For every } D_i^{\text{ind}}, \inf_{\text{stat}} D_i^{\text{ind}} = \text{Log} Z = N \text{Log} K.$$

Strictly speaking, this and the other values of \inf stated further on refer to the amount of information relative to some prior knowledge represented by a description D_0 of the physical system in question, say the gas g . Throughout the following discussion on information we regard D_0 as given. D_0 states the value of N , the total range R^*

¹ Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1949).

in the μ -space and a cell system Ω^μ with K cells. Thus we can infer from D_0 the volume V^μ of R^μ and the cell volume v^μ ; furthermore, in the γ -space, the total range R^γ and its volume V^γ , the cell system Ω^γ and the cell volume v^γ . [Note that D_0 does not state the energy E , and that (10-2) refers to the totality of all D^{ind} for the given N and K , not only to those of a given energy level.]

Together with Y. Bar-Hillel, I have developed a semantical theory of information.² In distinction to the statistical theory mentioned above, it takes into consideration the meaning or content of the statements. We have defined several concepts to explicate the amount of information. These concepts are, in our opinion, closer than the statistical concept to the meaning intended by scientists in pre-systematic talk when they say that one statement is stronger in content or carries more information than another. While the statistical concept of information is based on the statistical concept of probability (which means relative frequency in the long run), our theory uses a logical or inductive concept of probability. For our present discussion, we shall use that function m_D of initial probability (see §2) which has equal values for all state-descriptions, that is, in the present context, for all D^{ind} :

$$(10-3) \quad \text{For every } D_i^{\text{ind}}, m_D(D_i^{\text{ind}}) = \frac{1}{Z}.$$

Then we define the *amount of semantical information*:

$$(10-4) \quad \text{For any statement } D_i, \inf_{\text{sem}}(D_i) =_{\text{df}} -\log m_D(D_i).$$

Hence:

$$(10-5) \quad \text{For every } D_i^{\text{ind}}, \inf_{\text{sem}}(D_i^{\text{ind}}) = \log Z = N \log K.$$

(In our treatise of 1952, the function \inf_{sem} based on m_D is denoted by ' \inf_D .'

We see from (10-2) and (10-5) that, under the assumptions made here, the functions \inf_{stat} and \inf_{sem} agree in their basic values and hence also in their derivative values. Therefore we shall write

² R. Carnap and Y. Bar-Hillel, *An outline of the theory of semantic information*, Res. Lab. of Electronics, M.I.T. Report No. 247, 1952. A short survey is given in: Bar-Hillel and Carnap, "Semantic Information," *Brit. J. Phil. of Science*, 4 (1953), pp. 147-157; this is a paper read by Bar-Hillel at the Symposium on Applications of Communications Theory (London, September 1952), London, 1953, pp. 503-512.

henceforth simply ‘inf,’ which may be interpreted either way. Thus:

$$(10-6) \quad \text{For every } D_i^{\text{ind}}, \inf(D_i^{\text{ind}}) = \log Z = N \log K.$$

For a disjunction of m different D^{ind} , both pr and m_D are m/Z . Hence:

$$(10-7) \quad \text{If } D_i \text{ is a disjunction of } m D^{\text{ind}}, \inf(D_i) = \log \frac{Z}{m}.$$

A D^{st} is a disjunction of $z D^{\text{ind}}$; therefore:

$$(10-8) \quad \inf(D_i^{\text{st}}) = \log \frac{Z}{z(D_i^{\text{st}})} = N \log K - \log z(D_i^{\text{st}}).$$

If the recipient of information obtains a description D_j in addition to his prior knowledge represented by D_i , the increase in information is called the relative amount of information of D_j with respect to D_i , denoted by ‘ $\inf(D_j | D_i)$ ’:

$$(10-9) \quad \inf(D_j | D_i) = \inf(D_i \cdot D_j) - \inf(D_i).$$

If D_j is logically implied by D_i , the conjunction $D_i \cdot D_j$ is logically equivalent to D_i ; therefore in this case:

$$(10-10) \quad \inf(D_j | D_i) = \inf(D_j) - \inf(D_i).$$

Let D_i^{st} correspond to D_i^{ind} ; then D_i^{st} is logically implied by D_i^{ind} . Hence:

$$(10-11) \quad \inf(D_i^{\text{ind}} | D_i^{\text{st}}) = \inf(D_i^{\text{ind}}) - \inf(D_i^{\text{st}}) = \log z(D_i^{\text{st}}).$$

From (10-8), (3-23), and (4-7):

$$(10-12) \quad S^*(D_i^{\text{st}}) = -\inf(D_i^{\text{st}}),$$

$$(10-13) \quad S_B(D_i^{\text{st}}) \cong -k \ln 2 \left[\inf(D_i^{\text{st}}) - N \log \frac{V^\mu}{N} \right].$$

The result (10-12) says that for a D^{st} the entropy S^* is equal to the negative of the amount of information. According to (10-13), a similar relation holds for S_B (aside from an additive constant, which could be abolished by changing the zero point of the S_B -scale, and a constant factor, which could be abolished by changing the unit). Or, if the negative of entropy is called ‘negentropy’ as suggested by

Brillouin, the negentropy of a D^{st} is equal or proportional to its amount of information.

In view of this result some authors declare that negentropy is always the same as amount of information. Whether this is the case or not depends upon the method used for extending S to other than statistical descriptions. Let D_i be a disjunction of $m D^{\text{ind}}$ with respect to a given cell system, with N and K fixed. Then $\inf(D_i) = \log(Z/m)$ (10-7), $= -\log m + N \log K$. Hence from (6-11):

$$(10-14) \quad S_B^{\text{II}}(D_i) = -k \ln 2 \left[\inf(D_i) - N \log \frac{V^\mu}{N} \right].$$

We extended S^* to $S^{*\text{I}}$ in (3-31). $S^{*\text{II}}$ would be defined in analogy to (6-11) as follows. Let D_i be a disjunction of $m D^{\text{ind}}$ ($1 \leq m \leq Z$), then:

$$(10-15) \quad S^{*\text{II}}(D_i) =_{\text{Def}} \log \frac{m}{Z}.$$

Hence:

$$(10-16) \quad S^{*\text{II}}(D_i^{\text{ind}}) = -\log Z = -N \log K.$$

$$(10-17) \quad S^{*\text{II}}(D_i^{\text{st}}) = \log \frac{z(D_i^{\text{st}})}{Z} = S^*(D_i^{\text{st}}).$$

With (6-14):

$$(10-18) \quad S^{*\text{II}}(D_i) = \log p_B(D_i).$$

With (10-7):

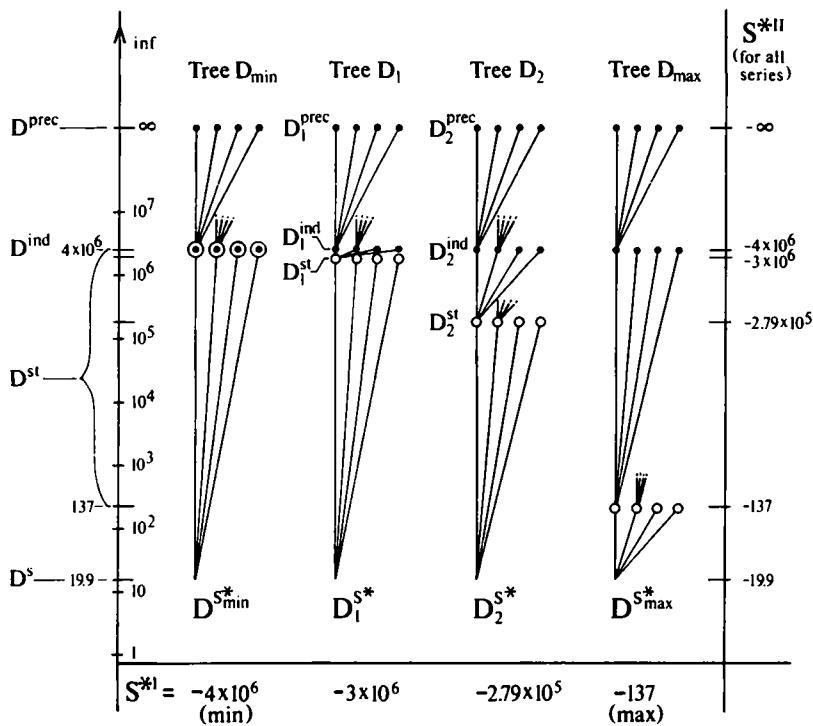
$$(10-19) \quad S^{*\text{II}}(D_i) = -\inf(D_i).$$

Thus for the S concepts constructed according to Method II, the results previously found for D^{st} ((10-12) and (10-13)) hold generally for all D_i of the kind described, which includes the D^{st} and the D^{ind} as special cases. The description D_1^S in Example 4 in §6, which states directly the (approximate) value of S_B for g at t_1 , is a special case of D_i with $m/Z = 10^{-6}$. Therefore the above results hold for D_1^S too.

We found earlier that, for any disjunction of D^{ind} , Gibbs's concept S_G^{II} has, aside from an additive constant, the same value as S_B^{II} (see (8-8)). Therefore the relation to \inf is for S_G^{II} essentially the same as for S_B^{II} . However, the situation is different in the case to

which Gibbs's theorem (8-14) refers. In this case, the description D_1 of g at t_2 is a disjunction of D^{ind} , and the description D_2 is derived from D_1 for the later time t_2 with the help of the laws of mechanics. According to Liouville's theorem (7-15), D_1 and D_2 have the same range volume, therefore the same probability and the same amount of information. On the other hand, according to Gibbs's theorem, S_G^{II} is greater for D_2 than for D_1 . Thus the above relation between S_G^{II} and \inf does not hold in this case (this is possible because D_2 is not a disjunction of D^{ind}). It is sometimes asserted that the increase of S_G^{II} in the course of time is due to the loss of information, since our prediction of the state of g at a future time t , based on the measurements made at t_1 , allegedly becomes more and more uncertain with increasing t . But this assertion is clearly erroneous. On the basis of the laws of mechanics, which are deterministic, not statistical laws, D_1 can be derived from D_2 just as well as D_2 from D_1 ; therefore both are equally strong, there is no decrease in the amount of information.

In order to clarify the situation with respect to Methods II and I, let us consider a series of descriptions D_1 of the following kind (for a given D_0 , i.e., fixed N and K). We represent the descriptions of the series by points on a vertical line (see fig. 3). We begin the series with an arbitrary precise description D_1^{prec} ; its range in the γ -space consists of one point U_1 . Let Q_1^γ be the γ -cell to which U_1 belongs. We take next some descriptions whose ranges include U_1 and are more and more comprehensive subclasses of Q_1^γ , down to that D^{ind} whose range is Q_1^γ itself; let this be D_1^{ind} . Then we proceed to more and more comprehensive disjunctions consisting of D_1^{ind} and other D^{ind} with the same μ -cell numbers, down to the disjunction of all these D^{ind} , which is the corresponding D_1^{st} . Then we take more and more comprehensive disjunctions of D_1^{st} and other D^{st} with the same value of z , say z_1 , down to the disjunction of all of them; let this be D_1^z . This corresponds to a point on the z -scale. Next we take a series of more and more comprehensive disjunctions of several D^z with z -values close to z_1 . This corresponds to more and more inclusive intervals on the z -scale around z_1 . Let this series end with a description stating the greatest interval around z_1 such that the proportion m_1/Z of all D^{ind} whose z -values belong to the interval does not exceed a fixed small fraction, say 10^{-6} . Let S be either S^* or S_B ; then S is a monotone increasing function of z . Therefore the last description states not only an approximate value of z but also an approximate value of S ; let this description be D_1^S (it is similar to the D_1^s in Example 4 in §6). Let S^I be the extension of S according to Method I. Then the value of S^I remains the same



(For the numerical values, see table (10-20).)

Figure 3. Trees of Descriptions.

throughout the series of descriptions from D_1^{prec} to D_1^S (down to D_1^*) exactly the same, but further on only approximately, with less and less precision). On the other hand, the probability pr_B increases from 0 to 10^{-6} , and therefore the amount of information inf decreases from ∞ to 19.9. Thus the simple relation of equality between S^I and $-\text{inf}$ (or S^I as a given linear function of inf) which we found for D^{st} , will hold for D_1^{st} but cannot hold for any other description of the series.

Let us illustrate the situation by a *numerical example*. We use the function S^* , for which the simplest relation (10-12) holds. The results for S_B would be analogous. [If we assume, to simplify the example, that $N = V^*$, then $S_B(D_i^{\text{st}}) = k \ln 2 \times S^*(D_i^{\text{st}}) = -k \ln 2 \times \text{inf}(D_i^{\text{st}})$.] We take $N = 10^6$, $K = 16$. As μ -cell numbers in the series D_1 we take $N_1 = N_2 = N/2$, all others 0. Then we consider another series D_2 of the same kind, but with the cell numbers as in the example D_{16} at the end of §3 ($N = 99,000$, etc.). In the subsequent table (10-20) we refer only to four descriptions from each series: D^{prec} , D^{ind} , D^{st} , and D^S ; the other descriptions have intermediate values.

(10-20) *Numerical example for two series of descriptions with $N = 10^6$, $K = 16$.*

	pr_B	inf	$S^{*\text{II}}$ for both series	$S^{*\text{I}}$ for D_1	$S^{*\text{I}}$ for D_2
$D_1^{\text{prec}}, D_2^{\text{prec}}$	0	∞	$-\infty$	-3,000,000	-279,000
$D_1^{\text{ind}}, D_2^{\text{ind}}$	16^{-10^6}	4,000,000	-4,000,000	-3,000,000	-279,000
D_1^{st}	8^{-10^6}	3,000,000	-3,000,000	-3,000,000	-279,000
D_2^{st}	6.9^{-10^6}	279,000	-279,000	-279,000	-279,000
D_1^S, D_2^S	10^{-6}	19.9	-19.9	-3,000,000	-279,000

In the *diagram* (fig. 3), the series D_1 is represented by one of the lines in the tree D_1 . And similarly the series D_2 in the tree D_2 . The case that all molecules belong to the same μ -cell is represented by the tree D_{\min} . In this case, z has its minimum value 1, and hence $S^*(D^{\text{st}})$ has its minimum. A D^{st} of this kind coincides with a D^{ind} . The other extreme case is characterized by equal μ -cell numbers; here, z and $S^*(D^{\text{st}})$ have their maxima (Tree D_{\max}). Two descriptions belonging to different series are represented in the diagram by points on the same horizontal level if they have the same pr_B and therefore the same inf . Thus, all D^{prec} are on the same level;

likewise, all D^{ind} and all D^{S^*} . On the other hand, the D^{st} (represented in the diagram by little circles) are on different levels; the highest possible level of a D^{st} is that in the tree D_{\min} , which is the level of the D^{ind} (these D^{st} are represented by circles with dots); the lowest level is that in the tree D_{\max} .

Concerning $S^{*\text{II}}$ (10-15) and $S^{*\text{I}}$ (3-31), the following holds generally for any series of the kind described; for the descriptions of our example we see it from the table. Since $S^{*\text{II}}$ is a function of pr_B , it has the same value for descriptions on the same level in different series. (Therefore in the diagram the values are given in a marginal column.) Since $S^{*\text{II}}$ is equal to $-\inf$ (10-19), it is monotone increasing in each series. In contrast, $S^{*\text{I}}$ has the same value throughout each series, but different values for different series with different z . (Therefore, in the diagram, the values are given in a marginal row.) This situation makes it clear that $S^{*\text{I}}$ is a physical property; as we know, it is a certain function of $\ln z$ which (aside from the unit chosen) is in agreement with thermodynamic entropy S_{th} . On the other hand, $S^{*\text{II}}$ cuts across S_{th} and represents merely a logical property, in particular the negative amount of information.

Perhaps the adherents of Method II will reply that S_B^{II} ought not to be regarded as a logical concept, because its various values for D^{st} always coincide with those of the original concept S_B , whose character as a physical concept is not questioned. However, the fact of this coincidence does not prove that S_B^{II} is not a logical concept. For any physical magnitude M with a continuous scale, we can in general find a class C^M of descriptions D^M of the following kind: (1) each description D_i^M ascribes to a physical system a value x_i of M with some precision, in other words, a small interval of length Δ_i around x_i ; (2) for any one of these descriptions D_i^M , its amount of information is proportional to the value x_i in question: $\inf(D_i^M) = Ax_i$, with a fixed constant A . Suppose that there is a probability density function δ over the scale of M ; then the probability that the value of M lies within a small interval of length Δ around x_i is $\Delta \cdot \delta(x_i)$. Now we define the class C^M as follows: it contains those descriptions D_i^M in which, for any x_i , $\Delta_i = 2^{-Ax_i}/\delta(x_i)$. Then $\text{pr}(D_i^M) = 2^{-Ax_i}$; hence, with (10-1), $\inf(D_i^M) = Ax_i$. For $A = -1$, we should have $\inf = -x_i$; this is the case with S^* for the class of the D^{st} (10-12). Thus, there is nothing mysterious about the fact (10-12) that for a certain class of descriptions there is a simple relation between the physical value stated and a certain logical measure of the description.

The main result of our discussions is that *the general statement of equality of entropy and negative amount of information can be*

maintained only if Method II is chosen. However, in this case the resulting concept S^{II} (in any of its versions) is not a physical but a logical concept. The customary use of the term 'entropy' for this concept is apt to lead to confusion. The simplest way of solving the terminological problem would be to use for S^{II} the term 'negative amount of information'; hence for $-S^{II}$ not 'negentropy' but 'amount of information.'

Although the general identification of entropy (as a physical concept) with the negative amount of information cannot be maintained, there are certainly important relations between these two concepts. So far we have discussed only one relation, the equality or proportionality in the case of D^* . Another important relation was first pointed out by Szilard.³ He analyzes the much discussed paradox of Maxwell's demon.⁴ Suppose that the vessel containing a gas is divided by a wall into two parts *A* and *B* with equal temperature. Maxwell imagines a little demon who is able to observe individual molecules of the gas and to operate a door covering a hole in the wall. He opens the door only when a molecule of high velocity in *A* or one of low velocity in *B* approaches. Thus, without expenditure of work, the temperature in *B* is continually increased and that in *A* decreased, in contradiction to the second law of thermodynamics. Szilard points out that the demon or an automatic apparatus, in order to ascertain the velocity of the approaching molecule a_i , has to make an experiment, say with light rays reflected by a_i . This experiment involved inevitably the dissipation of some energy and hence a corresponding increase in entropy for the total system consisting of the gas, the apparatus, the radiation, etc. This increase in entropy cancels out the decrease brought about by the transfer of a_i to the other part of the vessel. Thus the second law is not violated.

Szilard's idea has been further developed by later authors. In particular L. Brillouin, in several articles,⁵ has investigated the relation between negentropy and amount of information. He analyzes in detail some typical experiments designed to supply information of a certain kind. He shows that the increase ΔS in the total entropy of the physical system to be investigated together with

³ L. Szilard, "Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen," *Zeitschr. f. Physik* 53 (1929), 840–856.

⁴ J. C. Maxwell, *Theory of Heat* (London: Longmans, Green, 1871); p. 328.

⁵ L. Brillouin, (1) "Maxwell's demon cannot operate: Information and Entropy I," *J. Applied Physics*, 22 (1951), 334–337. (2) "Physical entropy and information II," *ibid.*, 338–343. (3) "The negentropy principle of information," *ibid.*, 24 (1953), 1152–1163.

the surrounding experimental arrangements can never be less than the increase Δinf in the amount of information; in other words, the efficiency ϵ of the experiment, defined as $\Delta \text{inf}/\Delta S$, is always ≤ 1 (or $\leq k \ln 2$, if Boltzmann's S_B is used and inf is defined by (10-1), compare (10-3)). These and other results of Brillouin's are certainly interesting and clarify the situation with respect to Maxwell's paradox in the direction first suggested by Szilard. However, when Brillouin proceeds to identify negentropy with amount of information, I cannot follow him any longer. He defines entropy by $k \ln P$, where P is the number of "possible structures" or "possible states" or "complexions"; thus he means presumably something like S_B^{II} (6-11) (omitting the additive constant, as is frequently done). At any rate, it seems from his discussions that he implicitly uses what we have called Method II. He does not seem to be aware that the definition of S which he uses (and which he ascribes to Boltzmann and Planck) makes S a logical rather than a physical concept.

