On the Approach to Statistical Equilibrium

ARNOLD J. F. SIEGERT
Physics Department, Northwestern University, Evanston, Illinois
(Received August 12, 1949)

We consider as a random process the distribution of a gas in momentum space as function of the time. The probability of changes of the distribution in infinitesimal time intervals is assumed to be given by the "Stosszahlansatz." For the Rayleigh model of a gas we derive the probability of a distribution as a function of the time. For the Boltzmann gas with microscopic reversibility we show that the probability of a distribution approaches the stationary value in the limit of infinite time.

IN the usual derivation of the H-theorem the probability description is abandoned when, for the purpose of the calculation, the collision probability postulate ("Stosszahlansatz") is used as giving an expression for the number of collisions. Only the final result of the calculation is again interpreted as a probability statement. It therefore seemed to us of interest to attempt a discussion of the approach to equilibrium adhering to the probability description throughout the calculation, and, by treating the "distribution" or "Z state" —i.e., the set $n(t)$ of occupation numbers $n_i(t)$ in momentum space cell i at time t—as a random process[1] to obtain the probability of a distribution as a function of the time, or at least to investigate whether this probability approaches a limit for infinite time. The "Stosszahlansatz" is used in its original significance as a postulate expressing the probability of a change of the distribution in an infinitesimal time interval in terms of the occupation numbers and a matrix α whose off diagonal elements are the probabilities of individual collisions.

In Section I we have derived a formal expression for the probability function and discussed its limit for $t\rightarrow\infty$. Limit theorems of a more general scope[2] are probably known, but are not needed for our purpose.

In Section II we consider the random process generated by the simplified "Stosszahlansatz" describing the Rayleigh model[3] of a gas, which consists of non-interacting independent molecules. Changes in the system are caused by collisions with the atoms of another gas which is not considered as part of the system and whose distribution is assumed to be unchanged throughout the process. The elementary transition probabilities for the model are then linear in the occupation numbers. The probability function for its distribution can be expressed explicitly if α can be diagonalized, which is shown to be the case, if e.g., the Rayleigh gas interacts with a gas in thermal equilibrium by binary collisions. The Stosszahlansatz of Section II also applies to the continuous time variant of a two-cell problem, which was suggested by Ehrenfest[4] as an entropy-model, further discussed by Schroedinger and Kohlrausch[5] and solved by Kac[6] (in its original form with discrete time variable) and later by the author[7] (with a continuous time variable). Another model with the simplified Stosszahlansatz is Ehrenfest's "wind tree" model.[8] The problem of concentration fluctuations[9] is also of this type, if n_i is interpreted as the number of colloidal particles in cell i in coordinate space, and if all intervals of observation times are long compared with the time it takes for the colloidal particles to come to Maxwell Boltzmann equilibrium in momentum space. This assumption makes it possible to consider the process as Markoffian in the occupation numbers in coordinate space cells (rather than in the occupation numbers in phase space). It was pointed out by M. Kac (private communication) that this process is actually (even with the above approximation) Markoffian in the whole set of occupation numbers, and that the occupation number of one cell is not (as is usually assumed) a Markoffian random process.

In Section III we have solved explicitly a simple two-cell problem with transition probabilities which are quadratic functions of the occupation numbers, as a model of a system with binary collisions.

In Section IV we have discussed the random process generated by the Boltzmann Stosszahlansatz with microscopic reversibility (symmetry of α). We have shown that the conditional probability of any distribution n, which can be reached from the initial distribution m by a succession of allowed binary collisions, approaches the multinomial coefficient for $t\rightarrow\infty$.

I

We consider a system described by a set of occupation numbers $n_1(t)$, $n_2(t)\cdots n_K(t)$, with

$$\sum_{i=1}^{K} n_i = N$$

[1] Terminology of M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945).
[2] Some far reaching limit theorems for random processes with a discrete time variable have been derived by R. v. Mises, *Wahrscheinlichkeitsrechnung* (Berlin, 1931).
[3] Lord Rayleigh, Scient. Papers III, p. 473, also M. C. Wang, Dissertation, Ann Arbor, Michigan, 1942.
[4] P. and T. Ehrenfest, Physik Zeits. **8**, 311 (1907); also Wang and Uhlenbeck, reference 1.
[5] E. Schroedinger and F. Kohlrausch, Physik Zeits. **27**, 306 (1926).
[6] M. Kac, Am. Math. Monthly **54**, 369 (1947).
[7] A. J. F. Siegert, LADC 438.
[8] P. and T. Ehrenfest, Encyclopaedie der Math. Wissenschaften **IV**, 32, p. 19; A. J. F. Siegert, Phys. Rev. **75**, 1322 (1949).
[9] M. v. Smoluchowski, Physik Zeits. **17**, 557 (1916); see also S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).

["distribution" or "Z state"] where $n_i(t)$ is the number of elements (e.g., atoms or molecules) in state i, where the state i may, e.g., be a cell in momentum space. The conditional probability

$$P(m/n, t) \equiv P(m_1, m_2 \cdots / n_1, n_2, \cdots; t)$$

is the probability that $n_i(t) = n_i$ if $n_i(0) = m_i$. We will try to calculate $P(m/n, t)$ from the probability of the change of the system in an infinitesimal time interval Δt: $P(m/n, \Delta t)$[10] assuming that

$$P(m/n, \Delta t) = (m|1|n) + (m|Q|n)\Delta t + o(\Delta t) \quad (1)$$

where $o(\Delta t)$ is defined by

$$\lim_{\Delta t \to 0} o(\Delta t)/\Delta t = 0$$

and where

$$(m|1|n) = \prod_i \delta_{m_i, n_i}.$$

This presupposes that $n(t)$ is K-dimensional Markoffian, i.e., that the probability of any change of the system in the time interval $(0, \Delta t)$ depends only on the occupation numbers at time zero. It is also assumed that the process is stationary, i.e., that $P(m/n, \Delta t)$ does not depend on the choice of the zero time. It is convenient to consider $P(m/n, t)$, $(m|Q|n)$ and $(m|1|n)$ as matrices $P(t)$, Q and 1 (unit matrix) with $2K$ indices which run from 0 to N. The matrix Q is arbitrary except for the properties:

$$(m|Q|n) = 0 \text{ unless } \sum_i m_i = \sum_i n_i = N \quad (2)$$

$$\sum_n (m|Q|n) = 0 \quad (3)$$

where $\sum_n$ is $\sum_{n_1 n_2 \cdots n_K}$ and

$$(m|Q|n) \geq 0 \text{ real}, \quad (4)$$

except if $n_i = m_i$ for all i. Q is independent of the time.

The above assumptions lead to two systems of differential equations:[11]

$$dP/dt = PQ, \quad (5)$$

$$dP/dt = QP \quad (6)$$

to be solved with the initial condition

$$P(0) = 1. \quad (7)$$

The formal solution is

$$P = \exp Qt. \quad (8)$$

[10] This change can be visualized as an elementary step of a random walk in the K-dimensional lattice formed by the points n on the hyperplane defined by $\Sigma_k n_k = N$.

[11] Special case of Eqs. (52) and (57), A. Kolmogoroff, Math. Ann. **104**, 415 (1931). The proof is based on the Smoluchowski equation

$$P(t+t') = P(t)P(t') \text{ for } t' \geq 0,$$

which implies that the validity of the above equations is restricted to cases where the state of the system can be ascertained between changes without disturbing the process.

The solution can be made explicit if Q can be diagonalized by means of a matrix C such that

$$(m|Q|n) = \sum_l (m|C^{-1}|l)\Lambda(l_1 l_2 \cdots l_K)(l|C|n). \quad (9)$$

If C can be found, we get

$$P(m/n, t) = \sum_l (m|C^{-1}|l)[\exp\Lambda(l)t] \cdot (l|C|n). \quad (10)$$

Because of Eq. (3), there is at least one right eigenvector to the eigenvalue $\Lambda(l_0) = 0$:

$$(n|C^{-1}|l_0) = b(l_0) \text{ (independent of } n).$$

If there is only one eigenvalue with vanishing real part we get, for $t \to \infty$

$$P(m/n_1 \infty) = b(l_0)(l_0|C|n) = (l_0|C|n)/\sum_{n'}(l_0|C|n') \quad (10')$$

since all $Re\Lambda(l) \leq 0$ (see Appendix I). In this case, therefore, an equilibrium probability function is approached, which is independent of the initial distribution m.

The eigenvalue zero of Q is degenerate, however, if e.g. the matrix Q is reducible by a mere rearrangement of rows and columns. Reducibility means that from a given initial point m in the K-dimensional lattice only the points n which belong to a sub-lattice can be reached, either directly or by any succession of steps. Reducibility can be caused by, e.g., conservation of energy and momentum. The above considerations are then applied to the sub-matrix and the sub-lattice, and if the eigenvalue zero of the sub-matrix is non-degenerate equilibrium is reached and is independent of the choice of the initial point m in the sublattice.

To apply these considerations to special cases, we will therefore have to prove in each case that Q can be diagonalized by a similarity transformation and that the eigenvalue zero of Q (or of its sub-matrices) is non-degenerate.

II

As a special case we consider a matrix Q representing changes in a system whose elements do not interact with each other, but change states independently, e.g., by collisions with a different set of molecules. This form of the "Stosszahlansatz" describes the Rayleigh model of a gas[3] and Ehrenfest's "wind-tree" model[8] (in momentum space), also the continuous variant of Ehrenfest's two box entropy model[7] and Smoluchowski's concentration fluctuations[9] (in coordinate space).

The matrix elements of Q are given by

$$(m|Q|n) = \sum_{\substack{i,k \\ i \neq k}}{}' \alpha_{ki} m_i \delta_{n_i, m_i - 1} \delta_{n_k, m_k + 1} \prod_{\substack{r \\ r \neq k, i}} \delta_{n_r, m_r} + \sum_i \alpha_{ii} m_i \prod_r \delta_{n_r, m_r}, \quad (11)$$

where $\alpha_{ki}\Delta t$ is the probability that a particle goes from

state i to state k in Δt. The off-diagonal matrix elements α_{ki} are arbitrary except for the conditions

$$\alpha_{ki} \geqslant 0 \text{ for } k \neq i \tag{12}$$

and the diagonal elements are defined by

$$\sum_k \alpha_{ki} = 0 \tag{13}$$

and it is assumed that α can be diagonalized by a matrix c such that

$$\alpha_{ki} = \sum_l c_{kl} \lambda_l c^{-1}{}_{li}. \tag{14}$$

We shall prove that under these assumptions Q is diagonalized by a certain matrix representation C of c, defined by Eqs. (24) and (25). To show this we define the generating polynomials

$$G_m(x, t) \equiv \sum_n P(m/n, t) \prod_j x_j{}^{n_j}, \tag{15}$$

$$Q_m(x) = \sum_n (m|Q|n) \prod_j x_j{}^{n_j}, \tag{16}$$

where x stands for $x_1, x_2, \cdots x_k$, and note that

$$Q_m(x) = \sum_{\substack{i,k \\ i \neq k}}{}' \alpha_{ki} m_i x_i{}^{m_i - 1} x_k{}^{m_k + 1} \prod_{\substack{j \\ j \neq k, i}} x_j{}^{m_j} + \sum_i \alpha_{ii} m_i \prod_j x_j{}^{m_j}$$

$$= \sum_{\substack{i,k \\ i \neq k}}{}' x_k \alpha_{ki} \frac{\partial}{\partial x_i} \prod_j x_j{}^{m_j} + \sum_i \alpha_{ii} x_i \frac{\partial}{\partial x_i} \prod_j x_j{}^{m_j} \tag{17}$$

or

$$Q_m(x) = \sum_{i,k} x_k \alpha_{ki} \frac{\partial}{\partial x_i} \prod_j x_j{}^{m_j}. \tag{18}$$

Equation (5) then yields

$$\frac{\partial G_m(x, t)}{\partial t} = \sum_{n'} P(m/n', t) \sum_n (n'|Q|n) \prod_j x_j{}^{n_j}$$

$$= \sum_{n'} P(m/n', t) \sum_{i,k} x_k \alpha_{ki} \frac{\partial}{\partial x_i} \prod_j x_j{}^{n_j} \tag{19}$$

$$= \sum_{i,k} x_k \alpha_{ki} \frac{\partial}{\partial x_i} G_m(x, t)$$

to be solved with initial conditions

$$G_m(x, 0) = \prod_j x_j{}^{m_j}. \tag{20}$$

It is convenient for our purposes to solve this equation by introducing a function

$$G(y, x, t) = N! \sum_m \prod_j \frac{y_j{}^{m_j}}{m_j!} G_m(x, t), \tag{21}$$

which satisfies the same differential equation as $G_m(x, t)$ with initial condition

$$G(y, x, 0) = (\sum_j y_j x_j)^N. \tag{22}$$

The solution is*

$$G(y, x, t) = [\sum_{i,j} x_i (e^{\alpha t})_{ij} y_j]^N$$

$$= [\sum_{ikj} x_i c_{ik} e^{\lambda_k t} c^{-1}{}_{kj} y_j]^N. \tag{23}$$

To obtain Q and P in the forms (9) and (10), respectively, we define a matrix representation $C(c)$ by

$$\prod_k u_k{}^{l_k} = \sum_n (l|C(c)|n) \prod_j x_j{}^{n_j} \tag{24}$$

with

$$u_k = \sum_i x_i c_{ik}. \tag{25}$$

We note that

$$[C(c)]^{-1} = C(c^{-1}), \tag{26}$$

$$\frac{N!}{\prod_k l_k!} \prod_k v_k{}^{l_k} = \sum_m \frac{N!}{\prod_r m_r!} \prod_r y_r{}^{m_r} (m|C(c^{-1})|l), \tag{27}$$

if

$$v_k = \sum_j c^{-1}{}_{kj} y_j. \tag{28}$$

We thus get

$$G(y, x, t) = [\sum_k u_k e^{\lambda_k t} v_k]^N$$

$$= \sum_l{}' \frac{N!}{\prod_k l_k!} \prod_k (u_k e^{\lambda_k t} v_k)^{l_k}$$

$$= \sum_l \frac{N!}{\prod_k l_k!} \prod_k u_k{}^{l_k} \exp(\sum_k \lambda_k l_k t) \prod_k v_k{}^{l_k}$$

$$= \sum_{m,n,l} \frac{N!}{\prod_r m_r!} \prod_r y_r{}^{m_r} (m|C(c^{-1})|l)$$

$$\times \exp(\sum_k \lambda_k l_k t) (l|C(c)|n) \prod_j x_j{}^{n_j}, \tag{29}$$

and using Eqs. (21), (15), and (26)

$$P(m/n, t) = \sum_l (m|C^{-1}|l) \exp(\sum_k \lambda_k l_k t) (l|C|n) \tag{30}$$

and

$$(m|Q|n) = \sum_l (m|C^{-1}|l) (\sum_k \lambda_k l_k) (l|C|n). \tag{31}$$

* See Appendix II for an alternative proof.

As a simple example we consider the most general two-state model, assuming

$$\alpha=\begin{pmatrix}-\gamma & \beta\\ \gamma & -\beta\end{pmatrix}. \tag{32}$$

The eigenvalues of α are $\lambda_1=0$ and $\lambda_2=-(\beta+\gamma)$ and the transformation matrix c is obtained as

$$c=\begin{pmatrix}\beta & 1\\ \gamma & -1\end{pmatrix} \text{ with } c^{-1}=\frac{1}{\beta+\gamma}\begin{pmatrix}1 & 1\\ \gamma & -\beta\end{pmatrix}. \tag{33}$$

Equation (23) becomes

$$\begin{aligned}G(y, x, t)&=(\beta+\gamma)^{-N}\{(\beta x_1+\gamma x_2)(y_1+y_2)\\ &\quad+(x_1-x_2)e^{-(\beta+\gamma)t}\\ &=(\beta+\gamma)^{-N}\{y_1[(\beta x_1+\gamma x_2)+\gamma(x_1-x_2)e^{-(\beta+\gamma)t}]\\ &\quad+y_2[(\beta x_1+\gamma x_2)-\beta(x_1-x_2)e^{-(\beta+\gamma)t}]\}^N\end{aligned} \tag{34}$$

and we get for $G_m(x, t)$:

$$\begin{aligned}G_m(x, t)&=(\beta+\gamma)^{-N}[(\beta x_1+\gamma x_2)+\gamma(x_1-x_2)e^{-(\beta+\gamma)t}]^{m_1}\\ &\quad\times[(\beta x_1+\gamma x_2)-\beta(x_1-x_2)e^{-(\beta+\gamma)t}]^{m_2}.\end{aligned} \tag{35}$$

For large times, $G_m(x, t)$ approaches

$$G_m(x, \infty)=\left[\frac{\beta x_1+\gamma x_2}{\beta+\gamma}\right]^N$$

and we get

$$P(m/n, \infty)=\frac{N!}{n_1!n_2!}\frac{\beta^{n_1}\gamma^{n_2}}{(\beta+\gamma)^N} \quad (\text{with } n_1+n_2=N). \tag{36}$$

The probabilities derived in references 4 and 5 can be obtained from Eq. (35).

We may interpret this example as describing a system of non-interacting spins in a constant magnetic field brought to thermal equilibrium, described by Eq. (36), by a temperature reservoir. The temperature T is defined by

$$\beta/\gamma=e^{(\epsilon_2-\epsilon_1)/kT},$$

where ϵ_1 and ϵ_2 are the energies of states 1 and 2, respectively, and k is the Boltzmann constant. The matrix α representing the effect of the temperature reservoir on the system must, therefore, be of the form

$$\alpha=\gamma\begin{pmatrix}-1 & e^{(\epsilon_2-\epsilon_1)/kT}\\ 1 & -e^{(\epsilon_2-\epsilon_1)/kT}\end{pmatrix}. \tag{37}$$

Just as in the case of Brownian motion,[12] the probability of changes of the system in Δt is thus derived from the validity of the Boltzmann law for the stationary distribution.

As a more general example we consider the case of a system which is brought to equilibrium by binary collisions with the atoms of a Boltzmann-gas. In this case α has the property

$$\alpha_{ki}=\alpha_{ik}e^{(\epsilon_i-\epsilon_k)/kT}, \tag{38}$$

where ϵ_i is the energy of state i, in addition to the properties expressed by Eqs. (12) and (13). To prove that α can be diagonalized we show that it can be symmetrized by a similarity transformation. We define the diagonal matrix ρ by

$$\rho_{ik}=e^{\epsilon_i/2kT}\delta_{ik} \tag{39}$$

and the symmetric real matrix σ by

$$\sigma=\rho\alpha\rho^{-1}. \tag{40}$$

σ can be diagonalized by a real matrix b, such that

$$\sigma=b\lambda b^{-1} \tag{41}$$

with λ real and diagonal, and with b real and $(b^{-1})_{il}=b_{li}$, thus

$$\alpha=\rho^{-1}\sigma\rho=\rho^{-1}b\lambda b^{-1}\rho=c^{-1}\lambda c, \tag{42}$$

where

$$c=b^{-1}\rho. \tag{43}$$

Next we show that the eigenvalue $\lambda_1=0$ is non-degenerate (unless α can be reduced by a mere re-arrangement of rows and columns) and derive the form of the eigen vector c_{il}. From

$$\lambda_l=\sum_{ik} b_{il}\sigma_{ik}b_{kl} \tag{46}$$

and

$$\sigma_{ii}=\alpha_{ii}=-\sum_{\substack{k\\ k\neq i}}{}' \alpha_{ki}=-\sum_{\substack{k\\ k\neq i}}{}' \rho_k^{-1}\sigma_{ki}\rho_i \tag{45}$$

we get, using the symmetry of σ,

$$\lambda_l=-\sum_{\substack{i,k\\ i\neq k}} b_{il}\rho_k^{-1}\sigma_{ik}\rho_i b_{il}+\sum_{\substack{i,k\\ i\neq k}} b_{il}\sigma_{ik}b_{kl} \tag{46}$$

and by interchanging indices

$$\lambda_l=-\sum_{\substack{i,k\\ i\neq k}} b_{kl}\rho_i^{-1}\sigma_{ik}\rho_k b_{kl}+\sum_{\substack{i,k\\ i\neq k}} b_{il}\sigma_{ik}b_{kl}. \tag{47}$$

Adding these expressions we get

$$\lambda_l=-\tfrac{1}{2}\sum_{\substack{i,k\\ i\neq k}}(b_{il}^2\rho_i\rho_k^{-1}-2b_{il}b_{kl}+b_{kl}^2\rho_i^{-1}\rho_k)\sigma_{ik} \tag{48}$$

$$=-\tfrac{1}{2}\sum_{\substack{i,k\\ i\neq k}}(b_{il}^2\rho_i^2-2b_{il}b_{kl}\rho_i\rho_k+b_{kl}^2\rho_k^2)\sigma_{ik}\rho_i^{-1}\rho_k^{-1} \tag{49}$$

$$=-\tfrac{1}{2}\sum_{\substack{i,k\\ i\neq k}}(b_{il}\rho_i-b_{kl}\rho_k)^2\sigma_{ik}(\rho_i\rho_k)^{-1}. \tag{50}$$

From this follows $\lambda_l\leqslant 0$, and if $\lambda_1=0$ $\rho_i b_{i1}=\rho_k b_{k1}$ unless $\sigma_{ik}=0$. Thus, if there exists at least one sequence of non-vanishing elements $\sigma_{ik_1}, \sigma_{k_1k_2}, \cdots\sigma_{k_{m-1}, k_m}, \sigma_{k_mk}$ between any pair ik we get $\rho_i b_{i1}=\rho_{k_1} b_{k_11}=\cdots=\rho_k b_{k1}$

[12] H. A. Kramers, Physica 7, 284 (1940).

or $b_{k1}=\rho_k^{-1}s$ where s is determined by

$$\sum_j b_{j1}^2=1=s^2\sum_j \rho_j^{-2}. \tag{51}$$

For the probability p_{ik} of an individual particle to go from state i to state k in time t we thus get

$$p_{ik}(t)\equiv(e^{\alpha t})_{ki}=(c^{-1}e^{\lambda}tc)_{ki}=\sum_j \rho_k^{-1}b_{kj}e^{\lambda_j t}b_{ji}^{-1}\rho_i \tag{52}$$

and for $t\to\infty$ only the term $j=1$ remains, so that

$$w_k\equiv p_{ik}(\infty)=\rho_k^{-1}b_{k1}b^{-1}{}_{1i}\rho_i=\rho_k^{-2}s^2 = e^{-\epsilon_k/kT}/\sum_j e^{-\epsilon_j/kT}. \tag{53}$$

Correspondingly one gets for the generating function (23)

$$G(y, x, t)\equiv(\sum_{i,k} x_i p_{ki}(t)y_k)^N \tag{54}$$

which becomes for $t\to\infty$

$$G(y, x, \infty)=(\sum_{i,k} x_i w_i y_k)^N=(\sum y_k)^N(\sum x_i w_i)^N \tag{55}$$

so that

$$P(m/n, \infty)=\frac{N!}{\prod_i n_i!}\prod_i w_i^{n_i} \tag{56}$$

which represents the canonical ensemble.

III

The cases discussed in Section II represent systems which change due to an outside influence. The elements of the system change states independent of each other, i.e. the system is actually an ensemble. If the system changes by interaction between the elements, $(n|Q|n')$ is no longer linear in the variables n_i, but can be, for instance, of second order (e.g. with Boltzmann's "Stosszahlansatz"). The solution for a simple model of this type is calculated. It is assumed that the elements can be in only two states, and that two elements in state 1 can interact and both change to state 2 and vice versa. The matrix Q is then given by

$$\begin{aligned}&(n_1n_2|Q|n_1'n_2')\\&=\beta\{n_1(n_1-1)\delta(n_1', n_1-2)\delta(n_2', n_2+2)\\&\quad+n_2(n_2-1)\delta(n_2', n_2-2)\delta(n_1', n_1+2)\\&\quad-[n_1(n_1-1)+n_2(n_2-1)]\delta(n_1', n_1)\delta(n_2', n_2)\},\end{aligned} \tag{57}$$

where β is a constant. It then follows that

$$\sum_{n_1'n_2'}'(n_1n_2|Q|n_1'n_2')x_1^{n_1'}x_2^{n_2'} =\beta(x_2^2-x_1^2)\left(\frac{\partial^2}{\partial x_1^2}-\frac{\partial^2}{\partial x_2^2}\right)x_1^{n_1}x_2^{n_2} \tag{58}$$

and, for the generating function, we get

$$\frac{\partial G_m(x, t)}{\partial t}=\beta(x_2^2-x_1^2)\left(\frac{\partial^2}{\partial x_1^2}-\frac{\partial^2}{\partial x_2^2}\right)G_m(x, t), \tag{59}$$

with initial condition

$$G_m(x, 0)=x_1^{m_1}x_2^{m_2}. \tag{60}$$

The solution is

$$G_m(x, t)=\sum_{\substack{l_1, l_2\\ l_1+l_2=N}}'(m|C^{-1}|l)e^{-4\beta l_1l_2t}(x_1+x_2)^{l_1} \times(x_1-x_2)^{l_2} \tag{61}$$

with $(m|C^{-1}|l)$ defined by

$$x_1^{m_1}x_2^{m_2}=\sum_{\substack{l_1, l_2\\ l_1+l_2=N}}'(m|C^{-1}|l)(x_1+x_2)^{l_1}(x_1-x_2)^{l_2} \tag{62}$$

in order to satisfy the initial condition. For $t\to\infty$, $G_m(x, t)$ approaches

$$G_m(x, \infty)=(m|C^{-1}|0, N)(x_1-x_2)^N +(m|C^{-1}|N, 0)(x_1+x_2)^N, \tag{63}$$

and since

$$(m|C^{-1}|N, 0)=2^{-N},\quad (m|C^{-1}|0, N)=2^{-N}(-)^{m_2}, \tag{64}$$

we get

$$G_m(x, \infty)=2^{-N}\{(-)^{m_2}(x_1-x_2)^N+(x_1+x_2)^N\} \tag{65}$$

and

$$P(m_1m_2/n_1n_2, \infty)=2^{-N}[1+(-)^{m_2+n_2}]\frac{N!}{n_1!n_2!} \tag{66}$$

with $m_1+m_2=n_1+n_2=N$. Except for the conservation of parity the binomial distribution is approached for $t\to\infty$.

IV

The problem of the approach to equilibrium of a gas with Boltzmann Stosszahlansatz—interpreted as a collision probability postulate—can be formulated in the same way. The matrix Q is given by

$$\begin{aligned}(m|Q|n)=\tfrac{1}{2}\{&\sum_{\substack{ijkl\\ ij\neq k, l}}\alpha_{kl, ij}m_im_j\delta_{n_i, m_i-1}\\&\times\delta_{n_jm_j-1}\delta_{n_k, m_k+1}\delta_{n_l, m_l+1}\prod_r\delta_{n_r, m_r}\\&+\sum_{ij}\alpha_{ij, ij}m_im_j\prod_r\delta_{n_rm_r}\},\end{aligned} \tag{67}$$

where $2\alpha_{kl, ij}\Delta t$, for $k, l\neq i, j$, is the probability of a collision $i, j\to k, l$ in Δt and $\alpha_{ij, ij}$ is defined by

$$\sum_{kl}\alpha_{kl, ij}=0 \tag{68}$$

$\alpha_{kl,ij}$ has the properties

$$\alpha_{kl,ii}=0, \tag{69}$$

$$\alpha_{kl,ij}=\alpha_{kl,ji}=\alpha_{lk,ij}, \tag{70}$$

$$\alpha_{kl,ij}\gtrless 0 \tag{71}$$

for $k,l\neq i,j$.

We also assume microscopic reversibility, i.e.,

$$\alpha_{kl,ij}=\alpha_{ij,kl}.^{13} \tag{72}$$

The matrix Q can be symmetrized by a similarity transformation

$$S=RQR^{-1}, \tag{73}$$

where R is defined by

$$(m|R|m')=(N!/\prod_r m_r!)^{\frac{1}{2}}\prod_r \delta_{m_r m_r'}. \tag{74}$$

To a submatrix of Q which can not be reduced by a mere interchange of rows and columns corresponds a similarly irreducible sub-matrix of S, since an element of S can be equal to zero only if the corresponding element of Q vanishes. Such a sub-matrix of Q defines a sublattice of distributions n such that each n can be reached from any other distribution m of the sub-lattice through a succession of steps mn', $n'n''\cdots n^{(s)}n$, for which the corresponding matrix elements $(m|Q|n')$, $(n'|Q|n'')\cdots(n^{(s)}|Q|n)$ are different from zero (i.e. through a sequence of allowed binary collisions). Equilibrium is reached in the sub-lattice according to Eq. (10′), since the eigenvalue zero is non-degenerate under these conditions. We get

$$P(m/n,\ \infty)=b_{l_0}(l_0|C|n) \tag{75}$$

for all distributions n which are in the same sub-lattice as m, where

$$(l_0|C|n)=(n|C^{-1}|l_0)(n|R^2|n)=b_{l_0}N!/\prod_r n_r! \tag{76}$$

and

$$b_{l_0}{}^2=[\sum_n{}' N!/\prod_r n_r!]^{-1} \tag{77}$$

the sum being extended over all distributions in the sub-lattice.

It would be of interest to see whether assumptions can be made concerning the matrix α such that the sub-lattice defined above becomes identical with the sub-lattice of all distributions with the same energy and momentum. The attempt to assign a discrete set of momenta and energies to the cells such that energy and momentum relations allow collisions between all pairs of particles lead, however, to number theoretical questions inherent in the discrete model which seem irrelevant to the physical problem.

[13] With these assumptions, the moments of $P(m/nt)$ satisfy a system of linear differential equations, see Appendix III.

APPENDIX I

In order that the right-hand side of Eq. (10) can be interpreted as a probability it is necessary that all $\Lambda(l)$ have non-positive real parts. This can be shown without using the assumption that Q can be diagonalized on the basis of Eqs. (3) and (4) only.

Suppose we have solved the system of equations

$$\sum_n(m|Q|n)(n|B|l)=(m|B|l)\Lambda(l),$$

which may be written in the form

$$-(m|B|l)\sum_{n\neq m}{}'(m|Q|n)+\sum_{n\neq m}{}'(m|Q|n)(n|B|l)=(m|B|l)\Lambda(l).$$

Now we chose among all $(m|B|l)$, for a chosen l, one, $(m_0|B|l)$ for which $|(m_0|B|l)|\geq|(n|B|l)|$ for all n. This may be assumed to be real without restricting generality. The equation for this index m_0 can be written as

$$\sum_{n\neq m_0}(m_0|Q|n)\left[\frac{(n|B|l)}{(m_0|B|l)}-1\right]=\Lambda(l).$$

Since $(m_0|B|l)\geq|(n|B|l)|\geq Re(n|B|l)$ we see that $Re\Lambda(l)\leq 0$ as required.

APPENDIX II

Equation (23) can also be derived using the independence of the particles: Let

$$(k_\sigma'|1|k_\sigma)+(k_\sigma'|\alpha|k_\sigma)\Delta t$$

denote the conditional probability that the particle numbered by σ and originally in cell k_σ, is after an infinitesimal time Δt in cell k_σ'. The conditional probability that particle σ is in cell k_σ' at t if it was in k_σ at $t=0$ is then given by the matrix

$$(k_\sigma'|e^{\alpha t}|k_\sigma)$$

and we get for the conditional probability $p(k_1k_2\cdots k_N/n,t)$ that $n_i(t)=n_i$, if the particles are at $t=0$ in the cells $k_1\cdots k_N$

$$p(k_1,k_2,\cdots k_N/n,t)=\sum_{k_1'\cdots k_N'}\prod_\sigma(k_\sigma'|e^{\alpha t}|k_\sigma)\prod_i(n_i|1|\sum_\sigma\delta_{i,k_\sigma'})$$

and for the generating function

$$\begin{aligned}g(k_1,k_2,\cdots/x,t)&\equiv\sum_n\prod_i x_i^{n_i}p(k_1\cdots k_N/n,t)\\&=\sum_{k_1'\cdots k_N'}\prod_\sigma(k_\sigma'|e^{\alpha t}|k_\sigma)\prod_i x_i^{\sum_\sigma\delta_{i,k_\sigma'}}\\&=\sum_{k_1'\cdots k_N'}\prod_\sigma\prod_i x_i^{\delta_{i,k_\sigma'}}(k_\sigma'|e^{\alpha t}|k_\sigma)\\&=\prod_\sigma(\sum_i x_i(i|e^{\alpha t}|k_\sigma))\\&=\prod_k[\sum_i x_i(i|e^{\alpha t}|k)]^{m_k}.\end{aligned}$$

This expression depends only on the number m_k of particles in cell k at time $t=0$ and not on the individual values k_σ. We have, therefore,

$$g(k_1\cdots k_N/x,t)=G_m(x,t)=\prod_k[\sum_i x_i(e^{\alpha t})_{ik}]^{m_k}$$

from which (22) follows using the definition (21). A still simpler derivation is obtained by introducing the random variables $n_k^{(\sigma)}$, with $n_k^{(\sigma)}=1$ if particle σ is in cell k, and zero otherwise. The generating functions for these variables are

$$P(0,0,\cdots,1,0\cdots/n^{(\sigma)},t)\prod_i x_i^{n_i{}^{(\sigma)}}=\sum_i x_i(e^{\alpha t})_{ik}$$

with the unity on the left-hand side in the kth place. Since $n_i=\sum_\sigma n_i^{(\sigma)}$, the above result for $G_m(x,t)$ follows from the independence of the particles.

APPENDIX III

The role of the non-linear Boltzmann equation in this formalism can be demonstrated by deriving the equation for the average $\bar{n}_i$.

We define the generating functions $Q_m(x)$ and $G_m(x, t)$ by Eqs. (15) and (16) and obtain

$$Q_m(x)=\tfrac{1}{2}\{\sum_{\substack{ijkl\\ij\neq kl}}\alpha_{kl;ij}m_im_jx_i^{m_i-1}x_j^{m_j-1}x_k^{m_k+1}x_l^{m_l+1}\prod_{r\neq ij,\,k,\,l}x_r^{m_r}$$

$$+\sum_{ij}\alpha_{ij,ij}m_im_j\prod_r x_r^{m_r}\}=\tfrac{1}{2}\sum_{ijkl}x_kx_l\alpha_{kl;ij}\frac{\partial^2}{\partial x_i\partial x_j}\prod_r x_r^{m_r}$$

and, therefore,

$$\frac{\partial G_m(x,t)}{\partial t}=\frac{1}{2}\sum_{ijkl}x_kx_l\alpha_{kl;ij}\frac{\partial^2}{\partial x_i\partial x_j}G_m(x,t).$$

For the averages

$$\bar{n}_s(t)=\left(\frac{\partial G_m}{\partial x_s}\right)_{x_1=x_2\cdots=x_K=1}$$

one obtains

$$\frac{d\bar{n}_s}{dt}=\sum_{ij,\,k}\alpha_{ks,\,ij}\langle n_in_j\rangle_{\mathrm{Av}}$$

using Eqs. (68) and (70). Since

$$\alpha_{ks,\,ks}=-\sum_{ij(ij\neq ks)}\alpha_{ij\,ks},$$

this can be written as

$$\frac{d\bar{n}_s}{dt}=\sum_{\substack{ijk\\ij\neq ks}}(\alpha_{ks,\,ij}\langle n_in_j\rangle_{\mathrm{Av}}-\alpha_{ij\,ks}\langle n_kn_s\rangle_{\mathrm{Av}}).^{14}$$

This becomes the Boltzmann equation

$$\frac{d\bar{n}_s}{dt}=\sum_{\substack{ijk\\ij\neq ks}}[\alpha_{ks,\,ij}\bar{n}_i\bar{n}_j-\alpha_{ij\,ks}\bar{n}_k\bar{n}_s]$$

if the correlations

$$\langle n_in_j\rangle_{\mathrm{Av}}-\bar{n}_i\bar{n}_j\equiv\langle(n_i-\bar{n}_i)(n_j-\bar{n}_j)\rangle_{\mathrm{Av}}$$

are neglected. Similar linear expressions can be derived for the time derivatives of higher moments. The non-linear Boltzmann equation is thus an approximation to the first one of a "hierarchy" of linear differential equations.

[14] This equation for the ensemble averages is thus in agreement with the transport equation for time averages derived by J. G. Kirkwood, J. Chem. Phys. **15**, 72 (1947), from essentially the same principles which underly the Stosszahlansatz.

PHYSICAL REVIEW VOLUME 76, NUMBER 11 DECEMBER 1, 1949

Structure of the Mercury Bands λ2540 and λ2476

S. MROZOWSKI
*Department of Physics, University of Chiacgo, Chicago, Illinois**

(Received August 22, 1949)

The structure of the band of Hg_2 at λ2540 and of the group of bands of Hg_2^+ at λ2476 was investigated under high dispersion. The band λ2540 was found to be limited by heads from both sides and to possess a number of sharp peaks on the long wave-length side. The bands at λ2476 were found to form a sequence with a gradually changing difference of the rotational constants of the upper and the lower energy states from negative to positive values. A number of peaks similar to those in λ2540 and band heads on both ends of bands are observed. An interpretation of the results is given in terms of the expected energy states of molecules of Hg_2 and Hg_2^+.

THE structure of the well-known mercury bands at λ2540 and λ2476 has been investigated under high dispersion. The photographs were taken several years ago and a short report on the results was made sometime later.[1] In this paper a more detailed account of the experiments will be given.

Both bands were investigated under very high resolution obtained in the second order of the 30-foot University of Chicago grating. The dispersion was about 0.44A per mm and the resolving power about 300,000. Unfortunately the mercury bands cannot be obtained under conditions which would give the full benefit of such a high resolving power. Due to temperature and pressure limitations the band λ2540 can be obtained in absorption with sufficient intensity only under conditions when the interesting finer details are wiped out to some extent by the temperature and pressure broadenings. In emission however the author succeeded in observing both bands under much more favorable conditions, and many fine details were obtained. The resolving power of the grating was still not utilized to the full extent (about the half of the resolving power would be sufficient to observe all details).

The band λ2540 was observed in absorption using a hydrogen discharge tube as a source of the continuum. The absorption tube (about 1 foot long) containing a few drops of mercury was sealed off under vacuum. A series of pictures was taken for different pressures of the saturated mercury vapor. The curve 3 in Fig. 1 gives the structure observed in absorption at temperatures below 200°C. At higher temperatures the band is more and more suppressed by the increasing continuous absorption extending from the resonance line λ2537 toward the longer wave-lengths and the structure becomes gradually wiped out. The gap in the absorption between λ2539.6 and λ2539.8 is very characteristic and was partly obtained by former investigators, as can be seen from a comparison of the curve 3 with the curves 1 and 2 in Fig. 1. Curve 1 was given in a paper by Wood

* Now at the Department of Physics, University of Buffalo, Buffalo, New York.

[1] S. Mrozowski, Phys. Rev. **73**, 1233 (1948).