A Test of the Linear Response Theory for Irreversible Processes

O. Seeberg
Institut für Theoretische Physik der Universität Frankfurt (Main)

Z. Naturforsch. **33a**, 110—120 (1978); received August 10, 1977

The temporal development of macroobservables is described within a correlation-function-formalism. The results are exact for a certain class of initial ensembles. The same problem is discussed with the help of the linear-response-formalism. The results agree under certain conditions which should be fulfilled for macroobservables.

Introduction

Van Kampen [1] has raised objections against the linear-response-theory in its applications to external disturbances. There is nothing to say against his arguments if we expect the linear-response-theory to be true for all initial ensembles and for all observables. In fact, the ordinary linear-response-theory always starts from the canonical equilibrium ensemble, though all observables are allowed. This question seems to be interesting with regard to the discussion whether quantum mechanics at all can be a suitable and complete basis for the description of macroscopic systems or not [2]. First one can argue that some additional principles must be introduced in order to describe macroscopic systems. Secondly it is possible that there is a new theory which contains quantum mechanics as a limiting case for very small systems, macrophysics on the other hand for very large systems. In this paper we describe another group of phenomena, the irreversible processes without external forces. It has been shown [3] that linear-response considerations equally well apply to these phenomena. On the other hand we can treat these phenomena in a completely other way without any use of linearization of the equations of motion. Moreover, this treatment can be done without any approximations, except for the choice of certain classes of initial ensembles: The initial ensembles must be purely macroscopic [4], [5]. Therefore it is possible to compare the results and thereby to give some criticism to the linear-response results. It turns out that the linear-response-formalism (l.r.f.) should be valid for macroobservables only. This result cannot be seen from a direct analysis of the l.r.f. It would be interesting to extend these considerations to the case of external disturbances [6]. This is not the aim of this paper.

Reprint requests to Doz. Dr. O. Seeberg, Robert-Mayer-Straße 8—10, D-6000 Frankfurt (Main).

In Sect. I we consider the case that the systems are completely isolated. In Sect. II we discuss the same phenomena for systems in contact with a heatbath. We remark that all interactions with the heatbath during the development of the macroobservables are disregarded, as well as in the l.r.f. with external forces. Why then do we start with the canonical ensemble? There is no physical reason for it.

Furthermore we give some remarks concerning the term

$$\tfrac{1}{2}\langle W^{\mathrm{eq}}(A_i A_j(t) + A_j(t) A_i)\rangle .$$

The time-correlation is precisely defined as the mean value of results of experiments on single systems. It turns out that this meaning cannot be given generally to the expression given above. But this is possible for a similar expression

$$\langle \tilde{W}^{\mathrm{eq}} A_i(t) A_j \rangle ,$$

where $\tilde{W}^{\mathrm{eq}} = C\, e^{-\beta \tilde{H}}$, $\tilde{H}$ the macroscopic energy (to be defined below), A_i, A_j macroobservables. Thus again a fully physical interpretation of the l.r.f. seems to be possible only for macroobservables, provided that both expressions agree very well. In order to check this assumption we must make some stability considerations, but this again is not the purpose of this paper.

In Sect. III we solve the problem with the help of l.r.f. techniques. In Sect. IV the results are compared and discussed. All mathematical details are discussed in the appendices.

1. Irreversible Processes in Closed Systems

We consider an ensemble $\mathfrak{G}$ of systems S with the number of particles N and volume V, described by a statistical operator W of the following form

$$W = P_{\mathfrak{r}_{E1}}\, W\, P_{\mathfrak{r}_{E1}} . \tag{1.1}$$

$P_{\mathfrak{r}_{E1}}$ is the projection-operator on to $\mathfrak{r}_{E_1}$, $\mathfrak{r}_{E_1}$ is the subspace in $\mathfrak{H}$, which is spanned by the eigenvectors φ_ν of the Hamiltonian H with eigenvalues $\varepsilon_\nu \in [E_1 - \Delta E, E_1[= J_{E_1}$. Let $\mathfrak{M}$ be the set of the macroobservables A_i. Without regarding the difficulties involved in the definition of the macroobservables we only demand the following properties [7], [8]

$$\begin{aligned} &\text{a)} \quad [A_i, A_j] = 0 && \text{for all } i, j, \\ &\text{b)} \quad A_i = \sum_k P_{\mathfrak{r}_{Ek}} A_i P_{\mathfrak{r}_{Ek}} && \text{for all } i. \end{aligned} \tag{1.2}$$

The property b) means: all macroobservables commute with the macroscopic energy $\tilde{H}$, $\tilde{H}$ being defined by

$$\tilde{H} = \sum_k P_{\mathfrak{r}_{Ek}} (E_k - \tfrac{1}{2} \Delta E) .$$

This is one possibility of defining $\tilde{H}$. Let $\mathfrak{r}_{E_1}$ be of finite dimension: $\dim \mathfrak{r}_{E_1} = d_1$. In the following section we shall write $\mathfrak{r}_{E_1} = \mathfrak{r}$, $d_1 = d$, for brevity. Now all linear operators on $\mathfrak{H}$ form a Hilbert-space Q, with the inner product $(A; B) = \mathrm{Sp}(A^+ B)$; [9].

The equation of motion for the statistical operator W now reads

$$\dot{W} = i \boldsymbol{L} W ; \quad \boldsymbol{L} B = BH - HB . \tag{1.3}$$

Now let $\mathfrak{B}$ be the set of the statistical operators in Q. We define a mapping $\boldsymbol{G}$ as follows [4].

$$\begin{aligned} &\boldsymbol{G}\colon \mathfrak{B} \to \mathfrak{B}, \\ &\boldsymbol{G} W = D \exp\{- \textstyle\sum \lambda_i A_i\} , \end{aligned} \tag{1.4}$$

with

$$\begin{aligned} &\text{a)} \quad \mathrm{Sp}((\boldsymbol{G}W) A_j) = \mathrm{Sp}(W A_j) , \\ &\text{b)} \quad \mathrm{Sp}(\boldsymbol{G}W) \quad\; = \mathrm{Sp}\, W = 1 . \end{aligned} \tag{1.5}$$

Conditions (1.5) determine λ_i and D (appendix A). The physical meaning of the mapping $\boldsymbol{G}$ is the following: $\boldsymbol{G}W$ is obtained from W by a variational principle:

$$\begin{aligned} &\delta\, \mathrm{Sp}(W \log W) = 0 \quad \text{with} \\ &\mathrm{Sp}(W A_i) = \langle A_i \rangle = \text{const.} \end{aligned}$$

Looking to $\mathrm{Sp}(W \log W)$ as to some kind of entropy, this variational principle corresponds in a certain sense to the second law. Now we are free to shift the scales in such a manner that

$$\langle A_i \rangle^{\mathrm{eq}} = \frac{1}{d} \mathrm{Sp}(P_\mathfrak{r} A_i) = 0 .$$

Then we get

$$\lambda_i^{\mathrm{eq}} = 0, \quad D^{\mathrm{eq}} = 1/d .$$

Note that the shift does not change the value of λ_i. Therefore we are led to the linearization of the operator $\boldsymbol{G}$, if we consider only statistical operators close to equilibrium with respect to the macroobservables A_i. This is a crucial point. We get from Eq. (1.4)

$$\boldsymbol{G} W \cong W^{\mathrm{eq}} + \sum D_i \lambda_i P_\mathfrak{r} - D^{\mathrm{eq}} \sum \lambda_i A_i .$$

Thus we can *define* a new mapping by

$$\boldsymbol{G_L} W = W^{\mathrm{eq}} + \sum D_i \lambda_i P_\mathfrak{r} - D^{\mathrm{eq}} \sum \lambda_i A_i \tag{1.6}$$

where D_i, λ_i again are to be determined by the conditions

$$\begin{aligned} &\mathrm{Sp}(\boldsymbol{G_L}(W) A_i) = \mathrm{Sp}(W A_i) = \langle A_i \rangle , \\ &\mathrm{Sp}(\boldsymbol{G_L} W) \qquad = 1 . \end{aligned} \tag{1.7}$$

From Eq. (1.7) we get with

$$A_i = P_\mathfrak{r} A_i , \quad W^{\mathrm{eq}} = D^{\mathrm{eq}} P_\mathfrak{r}$$

and $\langle A_i \rangle^{\mathrm{eq}} = 0$

$$\mathrm{Sp}(\textstyle\sum D_i \lambda_i P_\mathfrak{r}) = 0 \;\Rightarrow\; D_i = 0 .$$

Thus we get

$$\boldsymbol{G_L} W \quad = W^{\mathrm{eq}} - D^{\mathrm{eq}} \sum \lambda_i A_i .$$

Furthermore we have

$$\begin{aligned} &\boldsymbol{G_L} W^{\mathrm{eq}} = W^{\mathrm{eq}} - D^{\mathrm{eq}} \textstyle\sum \lambda_i^{\mathrm{eq}} A_i \quad \Rightarrow \\ &\langle A_j \rangle^{\mathrm{eq}} \quad = \langle A_j \rangle^{\mathrm{eq}} - \textstyle\sum \lambda_i^{\mathrm{eq}} \langle A_i A_j \rangle^{\mathrm{eq}} \quad \Rightarrow \\ &\textstyle\sum \langle A_i A_j \rangle^{\mathrm{eq}} \lambda_i^{\mathrm{eq}} = 0 . \end{aligned}$$

Now $\langle A_i A_j \rangle^{\mathrm{eq}}$ is a regular matrix (appendix B), thus $\lambda_i^{\mathrm{eq}} = 0$ and $\boldsymbol{G_L} W^{\mathrm{eq}} = W^{\mathrm{eq}}$. Furthermore we have from Eq. (1.7)

$$\mathrm{Sp}(W A_i) = - \sum \lambda_j \mathrm{Sp}(W^{\mathrm{eq}} A_j A_i) .$$

Defining γ_{ij} by

$$\sum \gamma_{ij} \mathrm{Sp}(W^{\mathrm{eq}} A_j A_k) = \delta_{ik} ,$$

we get

$$\begin{aligned} &\lambda_j = - \textstyle\sum \mathrm{Sp}(W A_i) \gamma_{ij} , \quad \text{or} \\ &\boldsymbol{G_L} W = W^{\mathrm{eq}} + W^{\mathrm{eq}} \textstyle\sum \mathrm{Sp}(W A_i) \gamma_{ij} A_j . \end{aligned} \tag{1.8}$$

Until now $\boldsymbol{G_L}$ is only defined on the set $\mathfrak{B}$ of the statistical operators. But we can define a new operator $\tilde{\boldsymbol{G}}$ on the Hilbert-space Q by

$$\tilde{\boldsymbol{G}} A = W^{\mathrm{eq}} \sum_{i,j} \mathrm{Sp}(A A_i) \gamma_{ij} A_j . \tag{1.9}$$

Thus we get

$$\boldsymbol{G_L} W = W^{\mathrm{eq}} + \tilde{\boldsymbol{G}} W , \quad \tilde{\boldsymbol{G}} W^{\mathrm{eq}} = 0 . \tag{1.10}$$

Now it can be shown that $\tilde{\boldsymbol{G}}$ is the projection-operator on to the subspace $\mathfrak{M}$ of the macro-observables spanned by the A_i (Appendix C). Then, under the *crucial* assumption [5], that

$$\tilde{\boldsymbol{G}}(W(0) - W^{\mathrm{eq}}) = W(0) - W^{\mathrm{eq}} \tag{1.11}$$

we get

$$\begin{aligned}\langle A_i\rangle(t) &= \mathrm{Sp}(A_i(W(t) - W^{\mathrm{eq}}))\\ &= (A_i; e^{i\boldsymbol{L}t}(W(0) - W^{\mathrm{eq}})) \qquad (1.12)\\ &= (\tilde{\boldsymbol{G}}\, e^{-i\boldsymbol{L}t}\tilde{\boldsymbol{G}} A_i; W(0) - W^{\mathrm{eq}}),\end{aligned}$$

where we have used that $\tilde{\boldsymbol{G}}A_i = A_i$. Equation (1.11) can be written in the following form

$$\langle A_i\rangle(t) = \sum_{k,j}\langle A_i(t)\, A_j\rangle^{\mathrm{eq}}\, \gamma_{jk}\, \langle A_k\rangle(0), \tag{1.13}$$

with $A_i(t) = e^{-i\boldsymbol{L}t}A_i$. This is a well-known result which e.g. can be found in [4]. Note, that this result is only valid under the assumption (1.11)! What does this condition mean?

$$\begin{aligned}&\tilde{\boldsymbol{G}}(W(0) - W^{\mathrm{eq}}) = \tilde{\boldsymbol{G}}W(0) = W(0) - W^{\mathrm{eq}},\\ &\tilde{\boldsymbol{G}}W(0) = \boldsymbol{G_L}\, W(0) - W^{\mathrm{eq}} \Rightarrow\\ &\boldsymbol{G_L}W(0) = W(0). \qquad (1.14)\end{aligned}$$

This condition means that only a very restricted class of initial ensembles is allowed. It follows from general reasons that some restricting conditions must be fulfilled. If there are no such conditions, the macroscopic equations never could be irreversible. A more detailed discussion can be found in [5]. Now it is possible to show [10], that

$$\langle A_i(t)\, A_j\rangle^{\mathrm{eq}}$$

is exactly the time-correlation-function for A_i, A_j which is given experimentally by the following procedure. Take $S \in \mathfrak{G}^{\mathrm{eq}}$, measure A_j, wait a time t, measure A_i (again on S.), and take the mean value of the product of the values over all $S \in \mathfrak{G}^{\mathrm{eq}}$.

Note that this result is true for *all* observables, if $\mathfrak{G}^{\mathrm{eq}}$ is the microcanonical equilibrium-ensemble.

With regard to the aim of this paper we have to clarify the connection between this correlation-function and the corresponding *canonical* one.

2. Irreversible Processes in Systems within a Heatbath

Let us repeat the considerations of Sect. I for systems which are composed of two systems, the first being a heatbath the second one the system S under consideration. $\mathfrak{H}$ has to be replaced by $\mathfrak{H}_{(1)} \times \mathfrak{H}_{(2)}$, H by $H_1 \times 1 + 1 \times H_2 + H_{12}$, A_i by $1 \times A_i$. H_{12} cannot be zero, otherwise there is no reason for development of an equilibrium-ensemble described by

$$W^{\mathrm{eq}} = \frac{1}{\alpha} P_{\mathfrak{r}}.$$

$\mathfrak{r}$ again is spanned by the eigenvectors of H with eigenvalues

$$\gamma_\nu \in [E - \Delta E, E[= J.$$

Now let us *assume* that

$$H_{12} = P_{\mathfrak{r}} H_{12} P_{\mathfrak{r}} + (1 - P_{\mathfrak{r}}) H_{12}(1 - P_{\mathfrak{r}}). \tag{2.1}$$

Equation (2.1) is a necessary and sufficient condition to be fulfilled, if we demand that $\mathfrak{r}$ can be spanned by the eigenvectors $(\psi_\varrho \varphi_\mu)$ of

$$H_0 = H_1 \times 1 + 1 \times H_2,$$

which are in $\mathfrak{r}$. Without this condition the usual derivation of the canonical ensemble already becomes complicated. Let us consider the statistical operator $\boldsymbol{G}W$ defined in Eqs. (1.4) and (1.5). Now we take the "Verkürzung" of $\boldsymbol{G}W$ [11] on to $\mathfrak{H}_{(2)}$ defined by

$$(\boldsymbol{G}W)^{(2)}_{\varrho\nu} = \sum_\mu \langle \chi_\mu\, \omega_\varrho |\, \boldsymbol{G}W\, | \chi_\mu\, \omega_\nu\rangle, \tag{2.2}$$

where (χ_μ), (ω_ϱ) are c.o.s. in $\mathfrak{H}_{(1)}$, $\mathfrak{H}_{(2)}$ respectively. Now let us choose $(\chi_\mu) = (\psi_\mu)$, $(\omega_\varrho) = (\varphi_\varrho)$. Then we get

$$(\boldsymbol{G}W)^{(2)}_{\varrho\nu} = D\, \langle \varphi_\varrho |\, \exp\{-\textstyle\sum \lambda_i A_i\}\, | \varphi_\nu\rangle\, D_{\varrho\nu}, \tag{2.3}$$

$D_{\varrho\nu}$ is the dimension of the subspace $\mathfrak{r}_{\varrho\nu} \subset \mathfrak{H}_{(1)}$ which is spanned by the eigenvectors of H_1 with eigenvalues

$$\begin{aligned}\eta \in [&E - \Delta E - K_{\varrho\nu} - \varepsilon_\varrho, E - K_{\varrho\nu} - \varepsilon_\varrho[\, \cap\\ &[E - \Delta E - K_{\varrho\nu} - \varepsilon_\nu, E - K_{\varrho\nu} - \varepsilon_\nu[,\end{aligned}$$

where $K_{\varrho\nu} = \langle \psi_\nu\, \varphi_\varrho |\, H_{12}\, | \psi_\nu\, \varphi_\varrho\rangle$.

Now the heatbath B is much larger than the subsystem S, therefore we choose ΔE to be much larger than the corresponding scale-length $\Delta\varepsilon$ for S. Now let us remember that $[\tilde{H}_2, A_i] = 0$. Therefore $\langle \varphi_\varrho |\, \exp\{-\sum \lambda_i A_i\}\, | \varphi_\nu\rangle$ vanishes, if φ_ϱ, φ_ν belong to different subspaces $\mathfrak{r}_{E_i}$, $\mathfrak{r}_{E_j}$. Let us therefore consider pairs of vectors inside one fixed $\mathfrak{r}_{E_i}$. Now it seems to be reasonable to *assume* that $D_{\varrho\nu}$ does not change very much, if φ_ϱ, φ_ν run in $\mathfrak{r}_{E_i}$. Thus we formulate the following assumption:

In the following calculations it is allowed to replace $D_{\varrho\nu}$ by $D_{E_iE_i}$, $D_{\varrho\varrho}$, $D_{\nu\nu}$ respectively. (2.4)

Furthermore the results do not change if we neglect $K_{\varrho\nu}$ (note that $K_{\varrho\nu}$ has the same order of magnitude as the interaction).

Now we use the familiar derivation of the canonical statistical operator (see e.g. [12], [13]). We get under certain assumptions which are not to be discussed in this paper

$$D_{\nu\nu} \cong e^{-\beta\varepsilon_\nu}\, C(N, N_1, E, \beta)\,, \tag{2.5}$$

where N is the total number of particles, N_1 the number of particles in S and $\beta = 1/kT$.

Thus we get two different expressions for $\boldsymbol{G}W^{(2)}$ from Equation (2.2). We choose the energy eigenvectors (ψ_μ) as c.o.s. in $\mathfrak{H}_{(1)}$, the eigenvectors (Φ_ϱ) of the macroobservables as c.o.s. in $\mathfrak{H}_{(2)}$. Then we get with the help of assumption (2.4)

$$(\boldsymbol{G}W)^{(2)}_{\varrho\nu} \cong D\langle\Phi_\varrho|\exp\{-\textstyle\sum\lambda_i A_i\}\,|\Phi_\nu\rangle\, D_{E_iE_i}$$
$$\text{for } \Phi_\varrho, \Phi_\nu \in \mathfrak{r}_{E_i}\,,\ i \text{ running, or}$$
$$\boldsymbol{G}W^{(2)} \cong D\exp\{-\beta\tilde{H} - \textstyle\sum\lambda_i A_i\}\,. \tag{2.6}$$

On the other hand again choosing (ψ_μ) as c.o.s. in $\mathfrak{H}_{(1)}$ but now the eigenvectors (φ_ϱ) of H_2 as c.o.s. in $\mathfrak{H}_{(2)}$ we get

$$(\boldsymbol{G}W)^{(2)}_{\varrho\nu} \cong \frac{D}{2}(D_{\varrho\varrho} + D_{\nu\nu}) \cdot \langle\varphi_\varrho|\exp\{-\textstyle\sum\lambda_i A_i\}\,|\varphi_\nu\rangle\,,$$

or

$$\boldsymbol{G}W^{(2)} \cong \frac{D}{2}\{e^{-\beta H}, \exp[-\textstyle\sum\lambda_i A_i]\}\,. \tag{2.7}$$

The anticommutator is used in Eq. (2.7) to make $\boldsymbol{G}W^{(2)}$ selfadjoint. Note that it is not permissible to write

$$\boldsymbol{G}W^{(2)} \cong D\exp\{-\beta H_2 - \textstyle\sum\lambda_i A_i\}\,,$$

because $[H_2, A_i] \neq 0$. Let us discuss both possibilities. The procedure is quite similar to that given in section I, therefore the discussion can be made brief. In order to avoid too many subscripts, we always shall use the same symbols, if no confusion is possible. We list the different steps. We begin with possibility (2.6).

1. Define $\boldsymbol{G}$ in $\mathfrak{H}_{(2)}$ by
$$\boldsymbol{G}W = D\exp\{-\beta\tilde{H}_2 - \textstyle\sum\lambda_i A_i\}\,,$$
where D, λ_i are to be determined from the conditions (1.5).

2. Shift the scales of the macroobservables so, that with $\tilde{W}^{\text{eq}} = \tilde{D}^{\text{eq}}\, e^{-\beta\tilde{H}_2}$ we get $\mathrm{Sp}(\tilde{W}^{\text{eq}} A_i) = 0$. It follows that $\tilde{\lambda}_i^{\text{eq}} = 0$.

3. Linearize:
$$\boldsymbol{G}W \cong (\tilde{D}^{\text{eq}} + \textstyle\sum\lambda_i D_i) \cdot e^{-\beta\tilde{H}_2} - \tilde{D}^{\text{eq}}\, e^{-\beta\tilde{H}_2}\textstyle\sum\lambda_i A_i\,,$$
$$\tilde{H}_2 \triangleq \tilde{H}\,.$$

4. Define
$$\boldsymbol{G_L}W = -\tilde{D}^{\text{eq}}\, e^{-\beta\tilde{H}}\textstyle\sum\lambda_i A_i + \tilde{D}^{\text{eq}}\, e^{-\beta\tilde{H}} + \textstyle\sum\lambda_i D_i\, e^{-\beta\tilde{H}}\,, \tag{2.8}$$
with the conditions (1.5). With 2. it follows that $D_i = 0$. Furthermore it follows that
$$\boldsymbol{G_L}W = \tilde{W}^{\text{eq}} + \tilde{W}^{\text{eq}}\textstyle\sum\mathrm{Sp}(WA_i)\,\gamma_{ij}\, A_j\,,$$
where $\sum\gamma_{ij}\,\mathrm{Sp}(\tilde{W}^{\text{eq}} A_j A_k) = \delta_{ik}$.

5. $\boldsymbol{G}$, $\boldsymbol{G_L}$ are only defined on the convex set of the statistical operators, therefore we define
$$\tilde{\boldsymbol{G}}A = \tilde{W}^{\text{eq}}\sum_{i,j}\mathrm{Sp}(AA_i)\,\gamma_{ij}\, A_j$$
as a linear operator in the Hilbert-space Q. Then we get
$$\boldsymbol{G_L}W = \tilde{W}^{\text{eq}} + \tilde{\boldsymbol{G}}W.$$

6. It is easily seen that $\tilde{\boldsymbol{G}} = \tilde{\boldsymbol{G}}^2$, this follows from the fact that $\mathrm{Sp}((\tilde{\boldsymbol{G}}A)A_i) = \mathrm{Sp}(AA_i)$. But $\tilde{\boldsymbol{G}}$ is not selfadjoint. We have
$$\tilde{\boldsymbol{G}}^{(+)}B = \textstyle\sum\gamma_{ik}\,\mathrm{Sp}(A_k\tilde{W}^{\text{eq}}B)\, A_i$$
(Appendix C). Therefore $\tilde{\boldsymbol{G}}$ is not a projection-operator in contradiction to our former results.

7. *Demand*:
$$\tilde{\boldsymbol{G}}(W(0) - \tilde{W}^{\text{eq}}) = W(0) - \tilde{W}^{\text{eq}}\,, \tag{2.8}$$
or $\boldsymbol{G_L}W(0) = W(0)$.

8. $$\begin{aligned}\langle A_k\rangle(t) &= \mathrm{Sp}(A_k\, e^{i\boldsymbol{L}_2 t}\, W(0)) \\ &= (A_k;\, e^{i\boldsymbol{L}_2 t}\, W(0)) \\ &= (A_k;\, e^{i\boldsymbol{L}_2 t}(W(0) - \tilde{W}^{\text{eq}})) \\ &= (A_k;\, e^{i\boldsymbol{L}_2 t}\tilde{\boldsymbol{G}}(W(0) - \tilde{W}^{\text{eq}})) \\ &= (\tilde{\boldsymbol{G}}^{(+)}\, e^{-i\boldsymbol{L}_2 t}A_k;\, W(0) - \tilde{W}^{\text{eq}})\,,\end{aligned}$$
or after inserting
$$\begin{aligned}\langle A_k\rangle(t) &= \textstyle\sum\gamma_{ji}\,\mathrm{Sp}(\tilde{W}^{\text{eq}}A_j\, e^{-i\boldsymbol{L}_2 t}A_k)\langle A_i\rangle(0) \\ &= \textstyle\sum\gamma_{ji}\,\mathrm{Sp}(\tilde{W}^{\text{eq}}A_j\, A_k(t))\langle A_i\rangle(0)\,.\end{aligned} \tag{2.9}$$

Note that $A_k(t) = e^{-iL_2 t} A_k$. That means, the system S is not disturbed by the heatbath, H_{12} has been neglected. But why then do we make use of the canonical ensemble? A similar question occurs in the l.r.f.

Another remark: The definition of $\boldsymbol{G}$ in step 1 remains unchanged, if we replace the operators A_i by $B_i = A_i + \gamma_i 1$. Hence $\tilde{\lambda}_i^{\mathrm{eq}} = 0$ for every shift. On the other hand it should be noted that

$$\boldsymbol{G} W^{\mathrm{eq}} \neq \tilde{W}^{\mathrm{eq}}, \quad W^{\mathrm{eq}} = C\, e^{-\beta H_2}.$$

Now let us linearize with respect to $\boldsymbol{G} W^{\mathrm{eq}}$. Almost all considerations remain unchanged. Replace only $\tilde{W}^{\mathrm{eq}}$ by $\boldsymbol{G} W^{\mathrm{eq}}$ in the definition of $\boldsymbol{G_L}$, γ_{ij}, $\tilde{\boldsymbol{G}}$, $\tilde{\boldsymbol{G}}^{(+)}$. Shift the scales to get

$$\mathrm{Sp}(W^{\mathrm{eq}} A_k) = 0 .$$

Demand:

$$\tilde{\boldsymbol{G}}(W(0) - W^{\mathrm{eq}}) = W(0) - W^{\mathrm{eq}}, \quad \text{or}$$

$$W(0) = W^{\mathrm{eq}} + \tilde{\boldsymbol{G}} W(0) .$$

Then the result is

$$\langle A_k \rangle (t) = \sum \gamma_{il}\, \mathrm{Sp}(\boldsymbol{G_L}(W^{\mathrm{eq}}) A_l A_k(t)) \langle A_i \rangle (0) .$$

Now these initial conditions look very strange. The corresponding condition (2.8, 7) is much clearer and therefore we choose this possibility.

Now we turn to possibility (2.7).

1. Define $\boldsymbol{G}$ in $\mathfrak{H}_{(2)}$ by
$$\boldsymbol{G} W = \frac{D}{2} \{e^{-\beta H}, \exp[-\textstyle\sum \lambda_i A_i]\}$$
with the conditions (1.5). Thus $\boldsymbol{G} W^{\mathrm{eq}} = W^{\mathrm{eq}}$.

2. Shift so that $\mathrm{Sp}(W^{\mathrm{eq}} A_i) = 0$.

3. Linearize
$$\boldsymbol{G} W \simeq (D + \textstyle\sum \lambda_i D_i)\, e^{-\beta H} - \frac{1}{2} D \{e^{-\beta H}, \sum \lambda_i A_i\}.$$

4. Define $\boldsymbol{G_L} W$ corresponding to 3.
$$\boldsymbol{G_L} W = (D + \textstyle\sum \lambda_i D_i)\, e^{-\beta H} - \frac{1}{2} D \{e^{-\beta H}, \sum \lambda_i A_i\} .$$
With the conditions (1.5) it follows that
$$D_i = 0\,, \quad \lambda_i = -\textstyle\sum \gamma_{ij} \langle A_j \rangle\,, \tag{2.10}$$
where
$$\textstyle\sum \gamma_{ij}\, \mathrm{Sp}(W^{\mathrm{eq}} A_j A_k) = \delta_{ik}\,,$$
where we have used the cyclic invariance of the trace.

5. Define
$$\tilde{\boldsymbol{G}} A = \tfrac{1}{2} \textstyle\sum \gamma_{ik}\, \mathrm{Sp}(A A_k) \{W^{\mathrm{eq}}, A_i\} .$$
Then we get
$$\tilde{\boldsymbol{G}}^{(+)} B = \tfrac{1}{2} \textstyle\sum \gamma_{ik}\, \mathrm{Sp}(\{W^{\mathrm{eq}}, A_k\} B) A_i .$$

6. *Demand* $\tilde{\boldsymbol{G}}(W(0) - W^{\mathrm{eq}}) = W(0) - W^{\mathrm{eq}} \Rightarrow$
$$\boldsymbol{G_L} W(0) = W(0).$$

7. As above we get
$$\langle A_k \rangle (t) = \tfrac{1}{2} \textstyle\sum \gamma_{ij}\, \mathrm{Sp}(\{A_j, A_k(t)\} W^{\mathrm{eq}}) \langle A_i \rangle (0) . \tag{2.11}$$

Now, if the equations of motion for $\langle A_k \rangle$ are stable with respect to small disturbances of the initial ensemble, we can expect that the solutions (2.9) and (2.11) agree approximately. Note that this stability could conversely serve for defining macroobservables [14].

Let us discuss the meaning of the term

$$\mathrm{Sp}(\tilde{W}^{\mathrm{eq}} A_l A_k(t)) .$$

Let us determine the mean value $\overline{A_l A_k(t)}$ defined by the procedure of Section I. Let $\mathfrak{G}$ be an ensemble of systems S described by the statistical operator W. To each system S a real number $\gamma[S]$ is assigned in the following way: Measure on S the observable A_k with result $\alpha[S]$. Wait a time t and then measure again on S the observable A_k with result $\beta[S]$. Then put $\gamma[S] = \alpha[S]\, \beta[S]$. Take the mean value over all S. The result is $\overline{\gamma[S]} = \overline{A_l A_k(t)}$. Let us calculate this number. Let be

$$\mathfrak{G}_{\alpha_\nu} = \{S \mid S \in \mathfrak{G} \wedge \alpha[S] = \alpha_\nu\} .$$

$\mathfrak{G}_{\alpha_\nu}$ is described by the statistical operator

$$W_{\alpha_\nu} = \frac{1}{d_{\alpha_\nu}} P_{\alpha_\nu} .$$

Now the probability for finding a value β_μ after a time t within $\mathfrak{G}_{\alpha_\nu}$ is equal to

$$P[\beta_\mu | \alpha_\nu] = \mathrm{Sp}\left(U_t^* \frac{P_{\alpha_\nu}}{d_{\alpha_\nu}} U_t P_{\beta_\mu} \right),$$

where $U_t = e^{iH_2 t}$. Therefore the probability for finding a pair (α_ν, β_μ) within $\mathfrak{G}$ is

$$P[\alpha_\nu, \beta_\mu] = P[\alpha_\nu]\, P[\beta_\mu | \alpha_\nu] .$$

But now we have $P[\alpha_\nu] = \mathrm{Sp}(W P_{\alpha_\nu})$.

Therefore we get

$$\overline{A_l A_k(t)} = \sum \alpha_\nu \beta_\mu\, \mathrm{Sp}(W P_{\alpha_\nu})\, \mathrm{Sp}\left(U_t^* \frac{P_{\alpha_\nu}}{d_{\alpha_\nu}} U_t P_{\beta_\mu} \right). \tag{2.12}$$

Let us choose $W = \tilde{W}^{\text{eq}}$. Then we get

$$\begin{aligned}\overline{A_l A_k(t)} &= \sum_\nu \sum_{\Phi_\varkappa \in \mathfrak{r}_\nu} \langle \Phi_\varkappa | W | \Phi_\varkappa \rangle \operatorname{Sp}\left(U_t^* \frac{P_{\alpha_\nu}}{d_{\alpha_\nu}} U_t A_k \right) \alpha_\nu \\ &= \sum_\nu \sum_{\Phi_\varkappa \in \mathfrak{r}_\nu} C \exp\{-\beta E_\varkappa^\nu\} \operatorname{Sp}\left(U_t^* \frac{P_{\alpha_\nu}}{d_{\alpha_\nu}} U_t A_k \right) \alpha_\nu \\ &= \sum \frac{\alpha_\nu}{d_{\alpha_\nu}} \sum_{\Phi_\varkappa \in \mathfrak{r}_\nu} \sum_{\Phi_\lambda \in \mathfrak{r}_\nu} \exp\{-\beta E_\varkappa^\nu\} \\ &\quad \cdot \langle \Phi_\lambda | U_t A_k U_t^* | \Phi_\lambda \rangle .\end{aligned}$$

If there is no degeneracy of the macroobservables, we may conclude

$$\begin{aligned}\overline{A_l A_k(t)} &= C \sum_\nu e^{-\beta E_\nu} \langle \Phi_\nu | U_t A_k U_t^* | \Phi_\nu \rangle \alpha_\nu \\ &= \sum_\nu \langle \Phi_\nu | \tilde{W}^{\text{eq}} U_t A_k U_t^* | \Phi_\nu \rangle \alpha_\nu \\ &= \sum_\nu \operatorname{Sp}(P_{\Phi_\nu} \tilde{W}^{\text{eq}} U_t A_k U_t^*) \alpha_\nu \\ &= \sum_\nu \operatorname{Sp}(\tilde{W}^{\text{eq}} P_{\Phi_\nu} \alpha_\nu U_t A_k U_t^*) \\ &= \operatorname{Sp}(\tilde{W}^{\text{eq}} A_l A_k(t)) . \end{aligned} \tag{2.13}$$

If on the other hand all $\Phi_\varkappa \in \mathfrak{r}_\nu$ belong to the same eigenvalue of the macroscopic energy, then we can write

$$\begin{aligned}\overline{A_l A_k(t)} &= \sum C e^{-\beta E_\nu} \alpha_\nu \operatorname{Sp}(P_{\alpha_\nu} U_t A_k U_t^*) \\ &= \operatorname{Sp}(\tilde{W}^{\text{eq}} A_l A_k(t)) . \end{aligned} \tag{2.14}$$

The same result of course is obtained by choosing $W = \boldsymbol{G_L} W^{\text{eq}}$. Note that we have used that W *commutes* with the macroobservables. Furthermore we must remember that the exact Hamiltonian H_2 does not commute with the macroobservables. Only in the case that $[W, A] = [W, B] = 0$, we can give to the expression $\operatorname{Sp}(WAB(t))$ the physical meaning of a time-correlation. From all these considerations we get a preference for Equation (2.9). Let us summarize. Provided, we are considering only initial ensembles $W(0)$ with

$$W(0) = \tilde{W}^{\text{eq}} + \tilde{W}^{\text{eq}} \sum \operatorname{Sp}(W(0) A_i) \gamma_{ij} A_j ,$$

we get the following solution:

$$\langle A_k \rangle (t) = \sum \gamma_{ij} \operatorname{Sp}(\tilde{W}^{\text{eq}} A_j A_k(t)) \langle A_i \rangle (0) .$$

The term $\operatorname{Sp}(\tilde{W}^{\text{eq}} A_j A_k(t))$ can be interpreted as a time-correlation. The initial condition can be interpreted in the following way: Only purely macroscopic initial ensembles are allowed.

3. The Solution of the Problem with the Help of the Linear —Response —Formalism

In the following considerations we follow Martin [3]. We consider an external disturbance

$$H^{\text{ext}} = -\sum A_j a_j(t) . \tag{3.1}$$

The usual l.r.f. then yields [15], [16]

$$\langle A_i \rangle (t) = \sum \int \chi_{ij}(t - t')\, e^{-\varepsilon |t'|} a_j(t')\, dt', \tag{3.2}$$

where

$$\chi_{ij}(t) = i \langle [A_i(t), A_j] \rangle_c^{\text{eq}}\, \Theta(t) .$$

Here $[A, B]$ is the commutator and

$$\langle A \rangle_c^{\text{eq}} = \operatorname{Sp}(C e^{-\beta H} A) .$$

We now choose the disturbance

$$a_j(t) = \begin{cases} \bar{a}_j \mid t < 0 , \\ 0 \mid t \geqq 0 . \end{cases}$$

Thus we get

$$\begin{aligned}\langle A_i \rangle (t = 0) &= \sum \int_0^\infty \chi_{ij}(\tau)\, e^{-\varepsilon\tau}\, \bar{a}_j\, d\tau , \\ \langle A_i \rangle_c^{\text{eq}} &= 0 \end{aligned} \tag{3.3}$$

or

$$\langle A_i \rangle (t = 0) = \sum_j \bar{a}_j\, \tilde{\chi}_{ij}(s = \varepsilon) ,$$

$\tilde{\chi}_{ij}$ being the Laplace transformed χ_{ij}. With the assumption that this linear system can be solved uniquely, we get

$$\begin{aligned}\bar{a}_k &= \sum (\tilde{\chi}(\varepsilon))_{ki}^{-1} \langle A_i \rangle (0) , \\ \langle A_i \rangle (t) &= \sum \int \chi_{ij}(t - t') \\ &\quad \cdot \exp\{-\varepsilon |t'|\} (\tilde{\chi})_{jl}^{-1} \langle A_l \rangle (0)\, dt' . \end{aligned} \tag{3.4}$$

Now let us take the Laplace transformation of Equation (3.4). After a short calculation we get

$$\begin{aligned}&\langle \tilde{A}_i \rangle (s = -i\Omega + \eta) \\ &= \sum \tilde{\chi}_{ij}(s = -i\Omega + \eta) \frac{1}{\varepsilon + i\Omega - \eta} \\ &\quad \cdot \sum (\tilde{\chi})_{jl}^{-1} (s = \varepsilon) \langle A_l \rangle (0) . \end{aligned} \tag{3.5}$$

In the next section we shall compare the result (3.5) with our earlier results.

4. Comparison of the Results

The l.r.f. calculations within these considerations can be found e.g. in [15]. We define

$$\begin{aligned} F_{ij}(t) &= \tfrac{1}{2}\,\mathrm{Sp}(W^{\mathrm{eq}}\{A_i(t), A_j\})\,, \\ S_{ij}(t) &= \mathrm{Sp}(A_i(t)\,A_j\,W^{\mathrm{eq}})\,, \\ \zeta_{ij}(t) &= i\,(S_{ij}(t) - S_{ji}(-t))\,, \\ \chi_{ij}(t) &= \Theta(t)\,\zeta_{ij}(t)\,. \end{aligned} \tag{4.1}$$

With $[W^{\mathrm{eq}}, H_2] = 0$ and the cyclic invariance of the trace we have

$$\mathrm{Sp}(W^{\mathrm{eq}}\,A_j\,A_i(t)) = S_{ji}(-t)\,.$$

Thus

$$F_{ij}(t) = \tfrac{1}{2}(S_{ij}(t) + S_{ji}(-t))\,. \tag{4.2}$$

S_{ij} is analytic in $\mathfrak{B} = \{z \,|\, -\beta \leqq \mathrm{Im}\,z \leqq 0\}$ provided, the occuring traces converge absolutely. Thus $S_{ij}(t - i\beta) = S_{ji}(-t)$. Let us consider the analytic function

$$G(z) = (1/\sqrt{2\pi})\;S_{ij}(z)\,e^{i\omega z}\,.$$

Then we get

$$\int_{\mathbb{R}} G(z)\,\mathrm{d}z = \int_{\mathbb{R}-i\beta} G(z)\,\mathrm{d}z\,.$$

Thus

$$\begin{aligned} \hat{S}_{ij}(\omega) &= (2\hat{F}_{ij}(\omega) - \hat{S}_{ij}(\omega))\,e^{\beta\omega}\,, \\ \hat{S}_{ij}(\omega) &= 2\hat{F}_{ij}(\omega)(1 + e^{-\beta\omega})^{-1}\,, \\ \zeta_{ij}(\omega) &= \frac{2i(1 - e^{-\beta\omega})}{1 + e^{-\beta\omega}}\,\hat{F}_{ij}(\omega)\,, \end{aligned} \tag{4.3}$$

which is the fluctuation-dissipation-theorem. For brevity

$$\frac{1 - e^{-\beta\omega}}{1 + e^{-\beta\omega}} = R(\omega) = \frac{\beta\omega}{2} + R'(\omega)\,;$$

$$\begin{aligned} \hat{F}_{ij}(\omega) &= \frac{1}{\sqrt{2\pi}}\left(\int_0^\infty e^{i\omega t}\,F_{ij}(t)\,\mathrm{d}t + \int_{-\infty}^0 e^{i\omega t}\,F_{ij}(t)\,\mathrm{d}t\right) \\ &= \hat{F}_{ij}^{(1)}(\omega) + \hat{F}_{ij}^{(2)}(\omega)\,. \end{aligned}$$

$\hat{F}_{ij}^{(1)}$ is analytic in the upper halfplane, $\hat{F}_{ij}^{(2)}$ correspondingly in the lower halfplane.

With

$$\tilde{f}(s) = \frac{1}{\sqrt{2\pi}}\int \frac{1}{i\,\omega + s}\,\hat{f}(\omega)\,\mathrm{d}\omega \tag{4.4}$$

we get after some calculation

$$\begin{aligned} \tilde{\zeta}_{ij}(s) &= \tilde{\chi}_{ij}(s) \\ &= \frac{4\pi i}{\sqrt{2\pi}}\,\hat{F}_{ij}^{(1)}(i\,s)\;R(i\,s) \\ &\quad + \frac{8\pi i}{\sqrt{2\pi}\,\beta}\left(\sum_{\mu=0}^{\infty}\frac{\hat{F}_{ij}^{(1)}(z_\mu)}{z_\mu - i\,s} - \sum_{\mu=-1}^{-\infty}\frac{\hat{F}_{ij}^{(2)}(z_\mu)}{z_\mu - i\,s}\right), \end{aligned} \tag{4.5}$$

$$z_\mu = i(2\mu + 1)\,\pi/\beta\,.$$

Note, that

$$\begin{aligned} \hat{F}_{ij}^{(1)}(i\,s) &= \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-st}\,F_{ij}(t)\,\mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}}\,\tilde{F}_{ij}(s)\,. \end{aligned}$$

Thus

$$\begin{aligned} \tilde{F}_{ij}(s) = \frac{1}{R(i\,s)\,2i}\Bigg(&\tilde{\chi}_{ij}(s) \\ &- \frac{8\pi i}{\sqrt{2\pi}\,\beta}\left(\sum_{\mu=0}^{\infty}\frac{\hat{F}_{ij}^{(1)}(z_\mu)}{z_\mu - i\,s}\right) - \sum_{\mu=-1}^{-\infty}\frac{\hat{F}_{ij}^{(2)}(z_\mu)}{z_\mu - i\,s}\Bigg)\Bigg). \end{aligned} \tag{4.6}$$

From Eq. (4.4), it follows that

$$\begin{aligned} \lim_{s\to 0}\chi_{ij}(s) &= \sqrt{\frac{2}{\pi}}\int \hat{F}_{ij}(\omega)\,\frac{R(\omega)}{\omega}\,\mathrm{d}\omega \\ &= \frac{\beta}{\sqrt{2\pi}}\int \hat{F}_{ij}(\omega) + \sqrt{\frac{2}{\pi}}\int \hat{F}_{ij}(\omega)\,\frac{R'(\omega)}{\omega}\,\mathrm{d}\omega\,. \end{aligned}$$

Furthermore

$$F_{ij}(0) = \mathrm{Sp}(W^{\mathrm{eq}}\,A_i\,A_j)$$

and

$$\gamma_{ij} = (F)_{ij}^{-1}\,.$$

Thus

$$\gamma_{ij} = \left(\lim_{s\to 0}\frac{\tilde{\chi}(s)}{\beta} - \sqrt{\frac{2}{\pi}}\int \hat{F}(\omega)\,\frac{R'(\omega)}{\omega\beta}\,\mathrm{d}\omega\right)_{ij}^{-1}. \tag{4.7}$$

Now let us consider the Laplace-transformation of Eq. (2.11)

$$\langle\tilde{A}_k\rangle(s) = \sum \gamma_{ij}\,\tilde{F}_{jk}(s)\,\langle A_i\rangle(0)\,. \tag{4.8}$$

Now we use the following *assumptions*:

$$\tilde{F}_{ij}(s) = -\frac{1}{\beta s}\,\tilde{\chi}_{ij}(s)\,,$$

$$\gamma_{ij} = \left(\lim_{s\to 0}\frac{\chi(s)}{\beta}\right)_{ij}^{-1}. \tag{4.9}$$

Inserting these expressions into Eq. (4.8), we get

$$\langle\tilde{A}_k\rangle(s) = \sum\left(\lim_{s\to 0}\tilde{\chi}(s)\right)_{ij}^{-1}\frac{1}{i\,s}\,\tilde{\chi}_{jk}(s)\,\langle A_i\rangle(0)\,,$$

or

$$\begin{aligned} &\langle\tilde{A}_k\rangle(s = -\,i\,\Omega + \eta) \\ &= \sum\left(\lim_{s\to 0}\tilde{\chi}(s)\right)_{ij}^{-1}\frac{1}{i\,\Omega - \eta}\,\tilde{\chi}_{jk}(-\,i\,\Omega + \eta)\langle A_i\rangle(0) \end{aligned} \tag{4.10}$$

which agrees *completely* with Equation (3.5).

We had used two assumptions on the way to this result.

1. Equation (2.11) is valid. Whether this is true or not must be investigated with the help of stability considerations. Equation (2.11) must be tested with respect to Equation (2.9). These stability considerations shall be done in a later work.

2. Assumptions (4.9) are valid. Let us start from Equation (4.5). Taking the inverse Laplace-transformations, we get:

$$\zeta_{ij}(t) = \frac{8\pi}{\sqrt{2\pi}\,\beta} \sum_{\mu=-1}^{-\infty} \hat{F}_{ij}^{(2)}(z_\mu) \exp\left\{\frac{(2\mu+1)\pi t}{\beta}\right\} + \frac{8\pi}{\sqrt{2\pi}\,\beta} \sum_{\mu=-1}^{-\infty} \hat{F}_{ij}^{(1)}(z_\mu) \exp\left\{\frac{(2\mu+1)\pi t}{\beta}\right\} + \sqrt{\frac{2}{\pi}} \sum_{\varrho} R(i\, s_\varrho) \int_{C_\varrho} \hat{F}_{ij}^{(1)}(i(s-s_\varrho))\, ds\, e^{s_\varrho t}. \tag{4.11}$$

where $\hat{F}_{ij}^{(1)}$ is the analytic continuation of $\hat{F}_{ij}^{(1)}$ into the left half-plane, s_ϱ its poles. Now the s_ϱ govern the behaviour of $F_{ij}(t)$, which should be smooth for macroobservables, thus, if $|s_\varrho| \ll \pi/\beta$, we can neglect the first terms in (4.11) in choosing an appropriate time-scale. Note, that π/β becomes big for small β, i.e. high temperatures. The same argument applies to our second condition. Starting from Eq. (4.7), we can neglect the second term, if $\hat{F}_{ij}(\omega)$ decreases sufficiently rapidly ($F_{ij}(t)$ smooth!) with respect to the fact, that

$$\lim_{\omega\to 0} R'(\omega)/\omega = 0 .$$

Of course we cannot expect that these assumptions are true for all observables. If we confine ourselves to macroobservables which should change *slowly* in time, these assumptions seem to be reasonable. Thus we are led to the conclusion that the l.r.f. can be justified for irreversible processes in closed systems and in systems in contact with a heatbath under the condition that the initial ensembles are purely macroscopic. Note that these conditions followed from a linearization with respect to equilibrium.

The discussion of the response of systems to external disturbances shall be discussed in a forthcoming paper.

Appendix A

We look for the solution of the variational problem $\delta(\mathrm{Sp}\, W \log W) = 0$, defined on $\mathfrak{H}$, under the conditions $\langle A_i \rangle = \text{const}$, $\mathrm{Sp}\, W = 1$. First we consider the case that the variation is done with a fixed c.o.s. (φ_ν). We start from $W = \sum w_\nu P_{\varphi_\nu}$. Now $\mathrm{Sp}(W \log W) = \sum w_\nu \log w_\nu$. The usual calculation yields

$$\bar{W} = \sum_\nu \exp\{-1 - \lambda_0' - \sum \lambda_i \langle \nu | A_i | \nu \rangle\} P_{\varphi_\nu}, \tag{A.1}$$

where λ_0, λ_i are to be determined by the conditions (1.5). We must prove the uniqueness of the solutions λ_0, λ_i. This can be done with the help of the considerations of Chintchin [16]. Let us define

$$\exp\{\sum \langle A_j \rangle \lambda_j\} \sum_\nu \exp\{-\sum \lambda_j \alpha_j^\nu\} = \Phi_A(\Lambda), \tag{A.2}$$

where $A_0 = 1$, $\lambda_0 = 1 + \lambda_0'$, $\alpha_j^\nu = \langle \nu | A_j | \nu \rangle$. Furthermore we put $\log \Phi_A = \Psi_A$. We look for the extrema of Ψ_A. We get

$$\frac{\partial \Psi_A}{\partial \lambda_k} = \langle A \rangle_k - \frac{\sum_\nu \exp\{-\sum \lambda_j \alpha_j^\nu\}\, a_k^\nu}{\sum \exp\{-\sum \lambda_j \alpha_j^\nu\}}, \quad k \geqq 1 . \tag{A.3}$$

Thus we have: If Λ_1 is an extremum of Ψ_A, then Λ_1 is a solution of our problem. Now let be Λ_1, Λ_2 two different extrema. We consider the straight line through Λ_1, Λ_2

$$\lambda_j = \Lambda_{1j} + \gamma(\Lambda_{2j} - \Lambda_{1j}), \quad \gamma \in \mathbb{R} .$$

Now we define the function $\Psi_A(\gamma)$ by inserting these values of Λ into Equation (A.2). Now $\Psi_A(\gamma)$ is a convex function. This is true if

$$\Phi_A''\, \Phi_A - (\Phi_A')^2 \geqq 0 .$$

Now we have

$$J(\gamma) = \Phi_A''\, \Phi_A - (\Phi_A')^2 = \sum_{\nu,\mu} R_\nu R_\mu \sum_{k,l} [(\alpha_k^\nu - \langle A_k \rangle)(\alpha_l^\nu - \langle A_l \rangle)(\Lambda_{2k} - \Lambda_{1k})(\Lambda_{2l} - \Lambda_{1l}) - (\alpha_k^\nu - \langle A_k \rangle)(\alpha_l^\mu - \langle A_l \rangle)(\Lambda_{2k} - \Lambda_{1k})(\Lambda_{2l} - \Lambda_{1l})],$$

where

$$R_\nu = \exp\{-\sum_j (\Lambda_{1j} + \gamma(\Lambda_{2j} - \Lambda_{1j}))\, (\alpha_j^\nu - \langle A_j \rangle)\} > 0 .$$

Thus we get

$$J(\gamma) = \sum_{\nu,\mu} R_\nu R_\mu \left(\sum_k (\Lambda_{2k} - \Lambda_{1k})(\alpha_k^\nu - \alpha_k^\mu)\right)^2 \geqq 0 .$$

Furthermore we have

$$J(\gamma)=0 \quad \Leftrightarrow \quad \sum_k (\Lambda_{2k}-\Lambda_{1k})(\alpha_k{}^{\nu}-\alpha_k{}^{\mu}=0 \quad \text{for all } \nu, \mu \,.$$

Now we have

$$\begin{aligned} T_{\nu\mu} &\overset{\text{def}}{=} \sum (\Lambda_{2k}-\Lambda_{1k})(\alpha_k{}^{\nu}-\alpha_k{}^{\mu}) \\ &= \langle\varphi_\nu|\sum(\Lambda_{2k}-\Lambda_{1k})A_k\,|\varphi_\nu\rangle \\ &\quad -\langle\varphi_\mu|\sum(\Lambda_{2k}-\Lambda_{1k})A_k\,|\varphi_\mu\rangle \,. \end{aligned} \tag{A 4}$$

We have to discriminate two different cases:

1. For the fixed c.o.s. (φ_ν) there is a linear combination $\sum(\Lambda_{2k}-\Lambda_{1k})A_k=B$ of the macroobservables with $T_{\nu\mu}=0$ for all ν, μ.

2. This is not the case.

In the case 2 we have:

$$T_{\nu\mu}=0 \quad \text{for all } \nu, \mu \quad \Rightarrow$$
$$\sum(\Lambda_{2k}-\Lambda_{1k})A_k=0$$

which implies, by means of the linear independence of the macroobservables, $\Lambda_{2k}-\Lambda_{1k}=0$. In this case the solution is uniquely determined, if it exists.

For the first case we do not answer the question concerning existence and uniqueness, because we do not need this answer as follows from later considerations.

Returning to the second case we prove the existence of the solution under the additional condition that all $w_\nu \neq 0$, or $\mathbb{T}w_\nu \neq 0$. The proof of the convexity of $\Psi_A(\gamma)$ shows that $\Phi_A(\Lambda)$ is convex for every straight line. Especially we consider the function

$$\begin{aligned} \Phi_A(\gamma) &= \sum_\nu \exp\{-\sum_j \gamma\, a_j(a_j{}^{\nu}-\langle A_j\rangle)\} \\ &= \sum e^{-\gamma s_\nu} \\ &= \sum_{S_\nu>0} e^{-\gamma s_\nu}+\sum_{S_\nu<0} e^{-\gamma s_\nu}+\sum_{S_\nu=0} e^{-\gamma s_\nu} \,. \end{aligned} \tag{A.5}$$

After presupposition not all s_ν can be zero. The non-vanishing cannot be all positive or all negative. We have

$$\begin{aligned} s_\nu &= \sum_j a_j(a_j{}^{\nu}-\langle A_j\rangle) \\ &= \langle\varphi_\nu|\sum a_j A_j\,|\varphi_\nu\rangle \\ &\quad -\sum w_\mu\langle\varphi_\mu|\sum a_j A_j\,|\varphi_\mu\rangle \end{aligned} \tag{A.6}$$

Let all non-vanishing s_ν be >0. Then we get

$$\begin{aligned} 0<\sum w_\nu s_\nu &= \sum w_\nu\langle\varphi_\nu|\sum a_j A_j\,|\varphi_\nu\rangle \\ &\quad -\sum w_\mu\langle\varphi_\mu|\sum a_j A_j\,|\varphi_\mu\rangle \\ &= 0 \,. \end{aligned}$$

Thus there must be positive and negative s_ν. But then we get

$$\Phi_A(\gamma) \geqq \exp\{-\gamma \operatorname*{Max}_{s_\nu>0} s_\nu\}+\exp\{-\gamma \operatorname*{Min}_{s_\nu<0} s_\nu\},$$

or

$$\begin{aligned} |\gamma| \geqq R \;\Rightarrow\; & \Phi_A(\gamma) \\ & \geqq \operatorname{Min}(\exp\{R \operatorname{Max} s_\nu\}, \exp\{R\,|\operatorname{Min} s_\nu|\}) \,. \end{aligned}$$

Thus there exists a sphere S of radius R with the following property

$$\exists \Lambda \in S \mid \Phi_A(\Lambda)<\Phi_A(\Lambda')$$

for all Λ' with $|\Lambda'|=R$.

Now $\Phi_A(\Lambda)$ is bounded and continuous, therefore it takes its minimal value inside S. Furthermore it is differentiable. This is the proof.

It can be shown by a very simple example that the condition $\mathbb{T}w_\mu \neq 0$ is necessary for the proof. Choose $A=\alpha P_{\varphi_1}+\beta P_{\varphi_2}|\alpha>\beta$ as operator in $\mathfrak{H}$ spanned by φ_1, φ_2. Choose $W=P_{\varphi_2}$. Then there exists no λ, D such that

$$\begin{aligned} &\operatorname{Sp}(D\,e^{-\lambda A})=1; \\ &\operatorname{Sp}(D\,e^{-\lambda A}A)=\operatorname{Sp}(WA) \,. \end{aligned} \tag{A.7}$$

In order to escape the restriction to a fixed c.o.s. we consider now the following situation. In any given ensemble $\mathfrak{G}$ we perform a measurement of the macroobservables A_i. The measured systems then form an ensemble $\mathfrak{G}'$ described by

$$W'=\sum u_\nu P_{\Phi_\nu} \,,$$

Φ_ν being the macroobservables,

$$u_\nu=\operatorname{Sp}(WP_{\Phi_\nu})=\operatorname{Sp}(\sum_\varrho w_\varrho P_{\varphi_\varrho} P_{\Phi_\nu}) \,.$$

But now we have

$$\sum_\nu u_\nu \log u_\nu \leqq \sum w_\nu \log w_\nu \,. \tag{A.8}$$

This follows from the inequality

$$(\sum q_i x_i)\log(\sum q_i x_i) \leqq \sum q_i x_i \log x_i$$

with

$$\sum q_i = 1$$

(see [17]). The equality only holds if all x_i are equal. Thus we have

$$\sum u_\nu \log u_\nu < \sum w_\nu \log w_\nu$$

unless all the w_ν are equal, but then

$$W=P_{\mathfrak{r}}\frac{1}{d}=W^{\mathrm{eq}} \,,$$

which is a trivial case. With

$$u_\nu = \sum_\varrho w_\varrho \gamma_{\varrho\nu}, \quad \gamma_{\varrho\nu} = |\langle\varphi_\varrho|\Phi_\nu\rangle|^2$$

we put $q_i \triangleq \gamma_{\varrho\nu}$, ν fixed, $x_i \triangleq w_\varrho$. Thus we get

$$u_\nu \log u_\nu \leqq \sum_\varrho \gamma_{\varrho\nu} w_\varrho \log w_\varrho,$$

which after summation is the proof. Furthermore we have $\mathrm{Sp}(\bar{W} \log \bar{W}) \leqq \mathrm{Sp}(W \log W)$. This follows from the fact that

$$\frac{\mathrm{d}^2}{\mathrm{d}\varepsilon^2}\Big(\sum_\nu (\bar{W}_\nu + \varepsilon v_\nu) \log(\bar{W}_\nu + \varepsilon v_\nu)\Big)(\varepsilon = 0) > 0$$

under the conditions $\sum v_\nu = 0$, $\sum \langle\nu|A|\nu\rangle v_\nu = 0$. Before we start our final considerations we prove the following lemma:

If the c.o.s. (φ_ν) is the c.o.s. of the macroobservables we get the following result. There is no linear combination B of the macroobservables with the property: All $\langle\varphi_\nu|B|\varphi_\nu\rangle = \langle\Phi_\nu|B|\Phi_\nu\rangle$ are equal. That means we have the case 2.

Proof: $\langle\Phi_\nu|\sum a_j A_j|\Phi_\nu\rangle = \langle\Phi_\nu|B|\Phi_\nu\rangle = b_\nu$.

All b_ν are equal $\Rightarrow B = b\,1 = b\,P_\mathfrak{r}$.

$P_\mathfrak{r}$ itself is a macroobservable and thus we had a linear relation between the macroobservables, which is not the case. Thus the variational problem is uniquely solvable if it is solvable at all. This is the case if $\mathbb{T}w_\nu \neq 0$. Now let $\hat{\mathfrak{G}}$ be the ensemble which originates from $\mathfrak{G}$ by measurement of the macroobservables, $\tilde{\mathfrak{G}}$ the ensemble which originates from $\mathfrak{G}$ by variation. Furthermore we define

$$\mathscr{H}[\mathfrak{G}] = \mathrm{Sp}(W \log W). \tag{A.9}$$

Let us consider the following diagram

$$\begin{tikzcd}
\mathfrak{G} \arrow[r] \arrow[d] & \hat{\mathfrak{G}} \arrow[dd] \\
\tilde{\mathfrak{G}} \arrow[r] & \hat{\tilde{\mathfrak{G}}} \\
 & \tilde{\hat{\tilde{\mathfrak{G}}}}
\end{tikzcd}$$

Let us consider only ensembles $\mathfrak{G}$ with $\mathbb{T}w_\nu \neq 0$. If the variational problem for $\mathfrak{G}$ is not solvable, we pass over to $\hat{\mathfrak{G}}$ with $\mathscr{H}[\mathfrak{G}] > \mathscr{H}[\hat{\mathfrak{G}}]$. Now it is easily seen that $\mathbb{T}\hat{w}_\nu \neq 0$ too, thus the variational problem is uniquely solvable, the solution is $\tilde{\hat{\mathfrak{G}}}$. Thus we have $\mathscr{H}[\mathfrak{G}] > \mathscr{H}[\tilde{\hat{\mathfrak{G}}}]$. If the variational problem for $\mathfrak{G}$ is solvable, there could be several solutions. If there is only one solution $\tilde{\mathfrak{G}}$, we have $\mathscr{H}[\mathfrak{G}] \geqq \mathscr{H}[\tilde{\mathfrak{G}}]$.

If $\tilde{\mathfrak{G}} = \tilde{\mathfrak{G}}^{\mathrm{eq}}$, we have $\tilde{\mathfrak{G}} = \hat{\tilde{\mathfrak{G}}} = \tilde{\hat{\tilde{\mathfrak{G}}}}$ and the only solution is $\mathfrak{G}^{\mathrm{eq}}$. If $\tilde{\mathfrak{G}} \neq \mathfrak{G}^{\mathrm{eq}}$, we have $\mathscr{H}[\tilde{\mathfrak{G}}] > \mathscr{H}[\hat{\tilde{\mathfrak{G}}}]$. Now the expectation values of the macroobservables coincide for $\mathfrak{G}$, $\hat{\mathfrak{G}}$, $\tilde{\mathfrak{G}}$, $\hat{\tilde{\mathfrak{G}}}$, $\tilde{\hat{\tilde{\mathfrak{G}}}}$. Thus we get $\mathscr{H}[\hat{\tilde{\mathfrak{G}}}] \geqq \mathscr{H}[\tilde{\hat{\tilde{\mathfrak{G}}}}]$, and $\tilde{\hat{\tilde{\mathfrak{G}}}}$ is the only solution. If there are several solutions the same considerations apply, again yielding $\tilde{\hat{\tilde{\mathfrak{G}}}}$ as the only solution. Thus we have the final result:

If only ensembles $\mathfrak{G}$ with $\mathbb{T}w_\nu \neq 0$ are regarded, the complete variational problem is uniquely solvable. The solution is given by $\tilde{\hat{\tilde{\mathfrak{G}}}}$.

The same considerations apply if we look for solutions of the equations

$$\mathrm{Sp}(D \exp\{-\beta \tilde{H} + \sum \lambda_i A_i\}) = \langle A_i\rangle.$$

The proof is quite analogous. We do not discuss here uniqueness and existence of the solutions of the equations

$$\mathrm{Sp}\left(\frac{D}{2}\,\{e^{-\beta H}, \exp[-\sum \lambda_i A_i]\}\,A_j\right) = \langle A_j\rangle.$$

There is no difficulty in the linearized case, and only this is discussed in the paper.

Appendix B

We prove the regularity of the matrix $\langle A_i A_j\rangle^{\mathrm{eq}}$.

$$\sum_j \mathrm{Sp}(W^{\mathrm{eq}} A_i A_j)\,x_j = 0$$

$$\text{and} \quad x_j = 0 \text{ not for all } i \Rightarrow$$

$$\sum_{i,j} \mathrm{Sp}(W^{\mathrm{eq}} x_i A_i x_j A_j) = 0 \Rightarrow$$

$$\mathrm{Sp}(W^{\mathrm{eq}}(\sum x_i A_i)(\sum x_i A_i)) = \mathrm{Sp}(W^{\mathrm{eq}} B^2, (B^2) = 0.$$

Now W^{eq} is $\geqq 0$ for every definition.

Therefore we can write

$$\begin{aligned}
\mathrm{Sp}(W^{\mathrm{eq}} B^2) &= \mathrm{Sp}(\sqrt{W^{\mathrm{eq}}}\sqrt{W^{\mathrm{eq}}}\sqrt{B^+}\sqrt{B^+}) \\
&= \mathrm{Sp}(\sqrt{B^+}\sqrt{W^{\mathrm{eq}}}\sqrt{W^{\mathrm{eq}}}\sqrt{B^+}) \\
&= (\sqrt{W^{\mathrm{eq}}}\sqrt{B^+}; \sqrt{W^{\mathrm{eq}}}\sqrt{B^+}).
\end{aligned}$$

Therefore we get

$$\sqrt{W^{\mathrm{eq}}}\sqrt{B^+} = 0 \Rightarrow \sqrt{B^+} = 0 \Rightarrow B = 0$$

and that is not the case, because the A_i are linear independent.

Appendix C

We investigate some projection-properties.

A) $\tilde{G}$, defined by

$$\tilde{G}A = W^{\mathrm{eq}} \sum \mathrm{Sp}(AA_i)\, \gamma_{ij}\, A_j\,,$$

$$W^{\mathrm{eq}} = \frac{1}{d} P_{\mathfrak{r}}\,,$$

is a projection-operator!

1. $\tilde{G}(\tilde{G}A) = W^{\mathrm{eq}} \sum \mathrm{Sp}(\tilde{G}AA_i)\, \gamma_{ij}\, A_j\,,$

$$\mathrm{Sp}((\tilde{G}A)\, A_i) = \sum \mathrm{Sp}(W^{\mathrm{eq}} A_j A_i)\, \mathrm{Sp}(AA_l) = \mathrm{Sp}(AA_i)\,,$$

thus $\tilde{G}^2 = \tilde{G}$, especially $\tilde{G}A_i = A_i$.

2. $\tilde{G}$ is selfadjoint!

$$\begin{aligned}(B; \tilde{G}A) &= (\tilde{G}A; B)^* \\ &= \mathrm{Sp}^*(\sum \mathrm{Sp}^*(AA_i)\, \gamma_{ij}\, A_j\, W^{\mathrm{eq}} B) \\ &= \sum \mathrm{Sp}(AA_i)\, \gamma_{ij}\, \mathrm{Sp}^*(A_j\, W^{\mathrm{eq}} B)\,. \\ (\tilde{G}B; A) &= \mathrm{Sp}(\sum \mathrm{Sp}^*(BA_i)\, \gamma_{ij}\, A_j\, W^{\mathrm{eq}} A)\,.\end{aligned}$$

Thus we have $(B; \tilde{G}A) = (\tilde{G}B; A)$.

B) $\tilde{G}$ defined in Eq. (2.8, 5) is not selfadjoint, the Hermitean conjugate is given by

$$\tilde{G}^{(+)} B = \sum \gamma_{ik}\, \mathrm{Sp}(A_k\, \tilde{W}^{\mathrm{eq}} B)\, A_i\,.$$

Proof:

$$\begin{aligned}(\tilde{G}^{(+)} B; A) &= \mathrm{Sp}(\sum \gamma_{ik}^*\, \mathrm{Sp}^*(A_k\, \tilde{W}^{\mathrm{eq}} B)\, A_i\, A) \\ &= \sum \gamma_{ik} (B; \tilde{W}^{\mathrm{eq}} A_k)(A_i; A) \\ &= (B; \sum \gamma_{ik}\, \mathrm{Sp}(AA_i)\, \tilde{W}^{\mathrm{eq}} A_k) \\ &= (B; \tilde{G}A)\,.\end{aligned}$$

C) $\tilde{G}$, defined in Eq. (2.10, 5) is not selfadjoint, the Hermitean conjugate is defined by

$$\tilde{G}^{(+)} B = \tfrac{1}{2} \sum \gamma_{ik}\, \mathrm{Sp}(\{W^{\mathrm{eq}}, A_k\}\, B)\, A_i\,.$$

Proof:

$$\begin{aligned}(\tilde{G}^{(+)} B; A) &= (B; \sum \gamma_{ik}\, \mathrm{Sp}(AA_i)\, \tfrac{1}{2} \{W^{\mathrm{eq}}, A_k\}) \\ &= (B; \tilde{G}A)\,.\end{aligned}$$

Appendix D

We prove the conjecture in Eq. (2.1)

$$\begin{aligned}&H_{12} = P_{\mathfrak{r}} H_{12} P_{\mathfrak{r}} + (1 - P_{\mathfrak{r}})\, H_{12} (1 - P_{\mathfrak{r}}) \;\Rightarrow \\ &[H_{12}, W^{\mathrm{eq}}] = 0, \quad (W^{\mathrm{eq}} = P_{\mathfrak{r}}/d) \;\Rightarrow \\ &[H_0, W^{\mathrm{eq}}] = 0 \;\Rightarrow\end{aligned}$$

$\mathfrak{r}$ can be spanned by some eigenvectors of H_0.

Conversion: Let a set φ_ϱ of eigenvectors of H_0 be a basis in $\mathfrak{r}$. Then we have

$$\begin{aligned}H_0 &= P_{\mathfrak{r}} H_0 P_{\mathfrak{r}} + (1 - P_{\mathfrak{r}})\, \mathrm{H}_0 (1 - P_{\mathfrak{r}}) \\ H &= P_{\mathfrak{r}} H P_{\mathfrak{r}} + (1 - P_{\mathfrak{r}})\, H (1 - P_{\mathfrak{r}}) \;\Rightarrow \\ H_{12} &= P_{\mathfrak{r}} H_{12} P_{\mathfrak{r}} + (1 - P_{\mathfrak{r}})\, H_{12} (1 - P_{\mathfrak{r}})\,.\end{aligned}$$

Acknowledgements

I whish to thank Professor Dr. R. Jelitto from the University of Frankfurt/Main for helpfull discussions in the course of a seminar and Professor Dr. W. Maass from the University of Marburg (Lahn) for many remarks.

[1] N. G. van Kampen, Physica Norvegica, vol. 5, Nos 3—4, 1971.

[2] G. Ludwig, Makroskopische Systeme und Quantenmechanik. Notes in Mathematical Physics, Marburg (Lahn), April 1972.

[3] P. C. Martin, Non Local Transport Coefficients Correlation Functions. In: Statistical Mechanics of Equilibrium and Non-Equilibrium, Ed. Meixner, North Holland, 1965.

[4] A. Katz, Principles of Statistical Mechanics, W. H. Freeman and Company 1967.

[5] G. Ludwig, Z. Phys. **173**, 232 (1963).

[6] W. M. Visscher, Phys. Rev. A, **10**, 6 (1974).

[7] G. Ludwig, Axiomatic Quantum Statistics of Macroscopic Systems (Ergodic Theory), Proc. of Int. School of Physics "Enrico Fermi", 1960.

[8] J. Schröter, Zur statistischen Mechanik, Endbericht für das Forschungsvorhaben Lu 29/9 der DFG.

[9] S. Grossmann, Funktionalanalysis, Akademische Verlagsgesellschaft Frankfurt/M. 1970.

[10] O. Seeberg, Ann. d. Phys., 7. Folge, **30**, 3/4 (1973).

[11] G. Ludwig, Die Grundlagen der Quantenmechanik, Springer-Verlag, Berlin 1954.

[12] R. Becker, Theorie der Wärme, Springer-Verlag, Berlin 1966.

[13] K. Huang, Statistische Mechanik, Bibliographisches Institut, Mannheim 1964.

[14] W. Maass, Private communication.

[15] R. Kubo, Linear Response Theory of Irreversible Processes. In: Statistical Mechanics of Equilibrium and Non-Equilibrium, Ed. Meixner, North-Holland, Amsterdam 1965.

[16] A. J. Chintchin, Mathematische Grundlagen der Statistischen Mechanik, Bibliographisches Institut, Mannheim 1964.

[17] G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities, Cambridge 1964.