The Case against Linear Response Theory

N. G. van KAMPEN
Institute for Theoretical Physics of the University, Utrecht, Netherlands

Kampen, N. G. van. The case against Linear Response Theory. *Physica Norvegica,* Vol. 5, Nos. 3–4, 1971.

The basic linearity assumption of linear response theory is shown to be completely unrealistic and incompatible with basic ideas of statistical mechanics of irreversible processes. (Editor's summary)

N. G. van Kampen, Institute of Theoretical Physics, Maliesingel 23, Utrecht, The Netherlands

DEAR HARALD:

I know that your knowledge of physics has deep roots, extending well into the days when science was still regarded as an intellectual adventure rather than an office job, when being a physicist was still a way of life rather than a way to make a living. This is the reason why I have hope that you may be willing to listen to some ideas that run counter to the current trend. For many years I have had doubts about the logic, if not the results, of linear response theory. Although I have met people who are inclined to agree, my arguments are not sufficiently concrete that I can expect to convince the unwilling. That is why I appreciate this opportunity to present them to you in private.

1. It goes without saying that macroscopic, nineteenth century physics abounds with phenomena in which some external force or agency acts on a system and provokes a response, which varies linearly with that force as long as it is not too strong. (It is true that this statement makes sense only if the scales in which force and response are measured are given, but this creates no difficulty because in most cases a natural scale exists.) Each linear dependence gives rise to a proportionality constant or 'phenomenological coefficient'. The task of statistical mechanics is not only to provide an expression for this coefficient in terms of molecular quantities, but also to understand how such a linear dependence arises from the microscopic equations of motion. In fact, it is hard to see how one could compute the coefficient without understanding the underlying mechanism.

Linear response theory does provide expressions for the phenomenological coefficients, but I assert that *it arrives at these expressions by a mathematical exercise, rather than by describing the actual mechanism which is responsible for the response.* To justify this assertion let me first outline the way in which linear response theory derives these expressions. It is convenient to make

---

The present work is dedicated to Professor Harald Wergeland on the occasion of his sixtieth birthday.

use of the quantum-mechanical formalism since it is more familiar than the classical treatment, although my discussion applies to classical mechanics as well.

2. Let a closed isolated system be subjected to an external force $F(t)$ such that its total Hamiltonian is

$$H(t) = H - F(t)Q\,,$$

where Q is some operator of the system. The density matrix ϱ of the system obeys

$$\dot{\varrho} = -i[H(t),\varrho] = -i[H,\varrho] + iF(t)\,[Q,\varrho]\,.$$

Equilibrium in the absence of F is described by

$$\varrho^{eq} = e^{-\beta H}/Z, \quad Z(\beta) = \mathrm{Tr}\, e^{-\beta H}.$$

The equation of motion can easily be solved to first order in F, assuming that F vanishes for $t \to -\infty$,

$$\varrho(t) = \varrho^{eq} + i\int_{-\infty}^{t} F(t')dt'[Q(t'-t),\varrho^{eq}]\,,$$

where the time-dependent operator $Q(t)$ is defined as usual:

$$Q(t) = e^{iHt}Qe^{-iHt}.$$

Having found $\varrho(t)$ one readily computes

$$\langle Q\rangle_t = \langle Q\rangle^{eq} + i\int_{-\infty}^{t} F(t')dt'\langle[Q(t),Q(t')]\rangle^{eq}.$$

This is a rigorous result for the expectation value of Q *to first order in F.* Note that it is not true that $\langle Q\rangle_t$ depends only on the value of F at the same t, as it involves the whole previous history of F as well. Accordingly, the 'proportionality constant' has become the kernel of a linear integral operator Γ,

$$\langle Q\rangle_t - \langle Q\rangle^{eq} = \int_0^{\infty}\Gamma(\tau)F(t-\tau)d\tau\,,$$

$$\Gamma(t) = i\langle[Q,Q(-\tau)]\rangle^{eq}.$$

In frequency language (with time factor $e^{-i\omega t}$) this simply means that the susceptibility

$$\chi(\omega) = i\int_0^{\infty}\langle[Q,Q(-\tau)]\rangle^{eq}e^{i\omega\tau}d\tau$$

is not just a constant, but a function of ω. This explicit expression enables one to verify the familiar formal properties of the susceptibility. For instance one readily verifies that $\chi(\omega)$ is holomorphic in the upper half of the complex ω-plane, and that $\chi(-\omega) = \chi^*(\omega)$ provided that there exists a representation in which both H and Q are represented by real kernels. On extending the treatment to the case of several observables Q_r, the latter property yields the reciprocity relation $\chi_{rs}(-\omega) = \chi^*_{sr}(\omega)$. Moreover, by comparing our expression for χ with a similar expression for the fluctuation spectrum one obtains the fluctuation-dissipation theorem.

3. Nonetheless, on closer examination the derivation outlined above appears to suffer from a fatal flaw, namely, the assumption that linearity of the macroscopic response means that the equation of motion for $\varrho(t)$ should also be solved to first order in the external field F. The equation for ϱ, however, is equivalent to the Schrödinger equation for the whole many-body system and contains therefore all the details of the microscopic motion of all individual particles. My point is that *linearity of the microscopic motion is entirely different from macroscopic linearity.* It is simply unrealistic to assume that the microscopic behaviour could be linear in F; the observed linearity of macroscopic phenomena is not based on it, but is a much more subtle statistical effect. Moreover I shall argue that microscopic linearity is not only unrealistic, but also incompatible with the very ideas on which the statistical mechanics of irreversible phenomena is based.

The gap between microscopic and macroscopic linearity is demonstrated by the following rough estimate of the range in which the microscopic motions are linear in F. Suppose the electrons in a conductor or semiconductor move freely, apart from occasional collisions with impurities. An external electric field F has the effect of shifting their paths; an electron that moves without collision during a time t will be shifted over a distance

$\frac{1}{2}t^2(eF/m)$. This affects the next collision, and in order that the effect is linear in F one must have

$$\frac{t^2}{2}\,\frac{eF}{m} \ll d\,,$$

where d is a measure for the diameter of impurity. If one generously takes $d \approx 100$ Å and allows a factor 100 for the effective mass, this inequality yields

$$t^2 F \ll 10^{-18}\ \text{sec}^2\ \text{Volt/cm}\,.$$

If one wants to apply the linearized solution of the equation for ϱ to a macroscopic measurement, t will be of the order of a second, so that F has to be less than 10^{-18} Volt/cm! (The exclamation mark is not superfluous, because when I mentioned this estimate at a conference, some people thought that the exponent was a misprint.)

Of course, in an ordinary conductor no electron moves freely during a second; however, the intervening collisions will have the effect of magnifying the distance between the paths in field F and in zero field, and hence even further reduce the range in which linearity holds.

This objection against linear response theory has sometimes been answered by pointing out that one need not trust the equation for $\varrho(t)$ itself – all that is needed is that the equation for $\langle Q\rangle_t$ should hold, which involves an average and is therefore no longer sensitive to microscopic details. However, the connection between $\varrho(0)$ and $\varrho(t)$ has been computed by means of the microscopic motions; once it is admitted that that is not permissible, the validity of the resulting formula for $\langle Q\rangle_t$ becomes a matter of wishful thinking.

4. The gap between microscopic and macroscopic linearity is so vast that using linear response theory cannot be regarded merely as stretching a point. Rather, it indicates that the whole approach is based on a fundamentally wrong picture of the actual mechanism involved. The following considerations are intended to show that the assumed *microscopic linearity is incompatible with the very basis of the statistical mechanics* of irreversible processes.

The miracle that statistical mechanics is called upon to explain, is the empirical fact that systems having 10^{23} degrees of freedom can be described, on a coarse, macroscopic scale by a much smaller number of variables, in such a way that these macroscopic variables again obey deterministic equations of motion. That is, their values at any given time uniquely determine their values at all later times. Nobody is able to specify precisely under which conditions this is true, but it is at least clear that the macroscopic variables are in some sense the slowly varying quantities of the system. The idea is that one can pick a time interval Δt, which on the one hand is so short that the macroscopic variables are practically constant; and on the other hand so large that in this time Δt the remaining microscopic variables of the system establish equilibrium, or rather a partial equilibrium with the values of the macroscopic variables as constraints. This state of partial equilibrium is uniquely specified by the values of the macroscopic variables alone and obeys the macroscopic equations. Thus *it is the randomization of the microscopic variables that enables one to eliminate them from the macroscopic picture.*

However, randomization implies that any information concerning the precise values of the microscopic variables is lost within a time interval Δt. A tiny uncertainty at the beginning of Δt has been magnified tremendously at the end of Δt. This magnification is incompatible with the assumption that an external field causes the microscopic variables to deviate linearly in F, since any initial small influence of F is also magnified. For this reason even the above estimate of 10^{-18} Volt/cm is still far too optimistic, because it ignores the randomizing effect of the intervening collisions, which is indispensable for the very existence of macroscopic laws.

To put it differently, statistical mechanics is concerned with the global aspects of the solutions of the equations of motion, those aspects that emerge as asymptotic properties after relatively long times. Linear response theory, however, treats the equations of motion in the way of textbooks on differential equations, which is appropriate for studying the detailed behaviour of a solution for sufficiently small times. Let us not forget that the breakthrough in understanding

Brownian motion came when Einstein and Smoluchowski realized that one does not see the actual path of the Brownian particle, but only the average displacement resulting from an extremely complicated microscopic motion. Linear response theory ignores this lesson.

You may well ask why this lack of randomization in linear response theory does not immediately lead to manifest absurdities. The reason is that *the effect of randomization is simulated by the linear approximation*. Equilibrium is assumed at $t=-\infty$, and the equations of motion are solved merely to first order in F; therefore the field always acts on the system as if it were in equilibrium. It is true that the expression for $\langle Q\rangle_t$ found in *2* involves the values of F at earlier times, but that only means that $\langle Q\rangle_t$ is the sum total of all previous responses. It does not contradict the fact that the way in which the system responds to a force F applied at t is independent of the previous history. My objection is that in linear response theory the previous history is not forgotten through randomization, but ignored through linearization.

5. The following model has helped me to visualize the state of affairs. Take the well-known Galton board, an early version of the pinball machine. When a pellet is dropped through the hole at the top, its horizontal coordinate x and its horizontal velocity v determine its microscopic state. Its vertical coordinate represents the time variable; for simplicity I shall regard the vertical velocity as constant. To turn it into a many-body system let us drop a large number of pellets successively, and consider the coordinate X of their center of gravity as the macroscopic variable Q. Unless all pellets have *precisely* the same initial x, the values of their x at the bottom will constitute a Gaussian distribution centered at the point $x=0$ vertically beneath the hole, so that $X=0$.

Now apply a horizontal force F, for instance by slightly tilting the board sideways. Clearly the center of the Gaussian will shift by an amount X that is linear in F, within some reasonable, macroscopic range. However, the individual paths of the pellets, and hence also their positions x

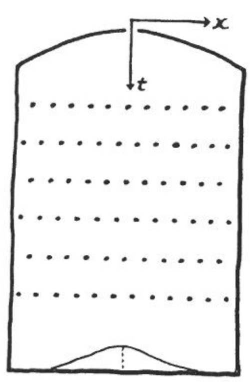

Fig. 1. The Galton board.

at the bottom, will change drastically. Clearly the linearity of the macroscopic response is not at all due to linearity of the microscopic motions. Rather the nonlinear deviations of the microscopic motions somehow combine to produce a linear macroscopic response.

Of course, for extremely small F the deviations in the individual x are linear, and this is the range to which linear response theory applies. An upper bound for this range is determined by the condition that the angle of tilting be small compared to the ratio d/T, where d is the diameter of a pin and T the total height of the board. (This necessary condition corresponds to the estimate in section *3*, and is still optimistic, inasmuch as the collisions with intervening pins further reduce the range of microscopic linearity.) In order to contrast the kinetic approach with linear response theory, I shall now apply it explicitly to this model.

6. Let $f(x,v,t)$ be the distribution of the pellets at a vertical distance t below the entrance. One assumes that it obeys a Boltzmann equation

$$\frac{\partial f}{\partial t}+v\frac{\partial f}{\partial x}+F\frac{\partial f}{\partial v}=\mathbf{W}f.$$

The operator $\mathbf{W}$ describes the collisions and is more explicitly

$$\mathbf{W}f(v)=\int\{W(v|v')f(v')-W(v'|v)f(v)\}dv',$$

with non-negative transition probabilities $W(v|v^1)$. The model is a kind of one-dimensional

Lorentz gas. It is convenient to reduce the problem by integrating over x so as to obtain an equation for the distribution function $g(v,t)$ of velocities alone,

$$\frac{\partial g}{\partial t} + F\frac{\partial g}{\partial v} = \mathbf{W} g\,.$$

Accordingly we shall measure the response to F by $\langle v\rangle$ rather than X; they are related by $X = \langle v\rangle T$.

The kinetic treatment is based on the idea that, barring a possible initial transient, a steady state is established, in which g no longer depends on t and can therefore be found from

$$F\frac{dg}{dv} = \mathbf{W} g\,.$$

It follows that the resulting average velocity

$$\langle v\rangle = \int vg(v)dv$$

depends on F alone. When $\langle v\rangle$ is expanded in powers of F the proportionality constant in the linear term is the phenomenological coefficient we are interested in.

In order to compute this coefficient suppose that $g_0(v)$ is the equilibrium distribution in the absence of F. Then substitute

$$g(v;F) = g_0(v) + Fg_1(v) + \ \ldots$$

in the equation and retain only the linear terms

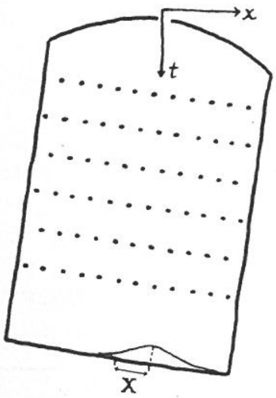

Fig. 2. An external force is applied.

$$\frac{dg_0(v)}{dv} = \mathbf{W} g_1(v)\,.$$

This integral equation for $g_1(v)$ can be solved, because the only summational invariant is a constant, to which the left-hand side is orthogonal. The solution is unique up to a term proportional to g_0, which is needed to satisfy the normalization condition for g but does not contribute to $\langle v\rangle$. Once the solution has been found the phenomenological coefficient is given by

$$\int vg_1(v)dv\,.$$

Incidentally, this procedure can be carried out explicitly for the somewhat artificial example $W(v|v') = p(v)q(v')$ with two positive functions p,q. If these functions are also even the resulting phenomenological coefficient is found to be

$$\tfrac{1}{2}\int \frac{p - vp'}{q^2}\,dv \Big/ \int \frac{p}{q}\,dv\,.$$

The exact nonlinear response can also be found in closed form for this example.

7. On comparing this kinetic treatment with the microscopic treatment in section *2* one observes three fundamental differences.

(i) *'Linearity' here means that one expands the distribution* φ, rather than the individual paths, to first order in F. The condition for validity is that the response $\langle v\rangle$ be small compared to the width of the equilibrium distribution $g_0(v)$. For conduction in metals and semiconductors this is amply satisfied. The diameter of the scatterers no longer enters into the estimate because the full dynamics of the scattering process is taken into account by means of $W(v|v')$.

(ii) *The time t has disappeared from the linear response*, owing to the assumption that g is identical with the steady state solution corresponding to the field F. (In case of time-dependent $F(t)$, deviations from this steady state will come in as corrections, giving rise to a frequency dependent phenomenological coefficient.) This means that the time-dependent equation for $g(v,t)$ needs to be solved only in the limit $t \to \infty$, so that the initial value of g is irrelevant. Consequently there is no need to rely on a force $F(t)$ that is switched on infinitely slowly.

(iii) *These desirable features have been obtained at the expense of a 'Stosszahlansatz'*: a randomness or molecular chaos assumption had to be made at all times (or at least at a sequence of closely spaced time points). As mentioned in *4*, linear response theory only needs a randomness assumption at the initial instant, because the assumed randomness remains operative for ever after, owing to the microscopic linearity.

These three differences show that linear response theory and the kinetic approach are radically different points of view. They are based on different and mutually exclusive concepts of the actual mechanism and are therefore incompatible. Only one can be right, the other is wrong. The arguments presented above convince me that linear response theory cannot be the right one.

Well, it is about time to rest my case and wait for your verdict. I have not discussed whether by some mathematical fluke the expressions that linear response theory finds for the transport coefficients happen to be correct. In case you are convinced by my arguments, that question is irrelevant; and if you are not, you will find that you cannot evaluate them anyway without using the same assumptions as kinetic theory. Nor did I discuss the question of higher orders beyond the linear approximation, because in my contribution to the book that Burgess edited under the title '*Fluctuations in Solids*' I have already argued that this extension of linear response theory is even more manifestly wrong. Finally you may observe that nowhere the thermodynamic limit is mentioned; the reason is that I feel that it has no bearing on the present argument.

Received 10 November 1971