The Road from Molecules to Onsager

N. G. van Kampen[1]

Our starting point consists of the *microscopic* dynamical equations of motion for the molecules, either classical or quantum mechanical. Subsequently the *repeated randomness assumption* is introduced, which breaks the time symmetry and produces the *mesoscopic* description in the form of a master equation for the probability distribution. Thereafter an *expansion in the reciprocal system size* leads to a *macroscopic* description, which may take one of two forms. Either it takes the form of a (nonlinear) deterministic rate equation for the macroscopic variables, tending to an equilibrium state; in this case the linearization around equilibrium produces the familiar Onsager reciprocal relations. Or it takes the form of a Fokker–Planck equation for the same variables; in that case a second expansion, this time in the temperature, leads to a nonlinear rate equation plus a dissipative term. The latter constitutes a nonlinear version of the Onsager equations.

1. INTRODUCTION

A thermodynamic system involves numerous degrees of freedom, be it the coordinates of molecules or the modes of some field. They obey *microscopic* equations of motion, but their motions are so rapid that the details are forever hidden to us. We know, however, that ultimately they lead to an equilibrium state. There is also a much smaller set of certain variables, which are functions of those same microscopic degrees of freedom, but their relative variation is much slower so that they can be followed by the macroscopic observer. These are the local densities of particles and energy etc.; they are the subject of fluid dynamics and of thermodynamics of

---

irreversible processes. The remarkable thing is that these *macroscopic* variables obey a closed set of equations of motion among themselves, not containing the other variables. The explanation is that, from the viewpoint of the rapid microscopic motion these slow variables are practically constant, so that the microscopic variables are able to adjust themselves to a partial equilibrium, that is, the equilibrium permitted by the constraints imposed by the instantaneous values of the macroscopic ones. Hence it is possible to eliminate the fast variables and establish equations for the evolution of the macrovariables A_r alone. These are the equations with which non-equilibrium thermodynamics is concerned.

When the whole process is confined to the vicinity of the over-all equilibrium these evolution equations will be linear and have the general form

$$\dot{A}_r = \sum_s M_{rs} A_s$$

Onsager[1] derived on general grounds a remarkable symmetry property of the coefficients M_{rs}. Much effort has been spent on generalizing this symmetry to the nonlinear regime[2] but it turned out that the inductive line of argument utilized by Onsager could not be extended to the nonlinear regime without running into ambiguities. For this reason it is necessary to find a more fundamental approach, starting from the very basis of molecular theory.

It is a long and arduous road from the microscopic equations of motion of the molecules to the macroscopic equations describing laboratory experiments. It involves all notorious problems of statistical mechanics. In particular we shall have to start from equations that are symmetric with respect to time reversal and end up with irreversible equations. Necessarily this requires some acrobatic steps that are less than rigorous. This cannot be avoided; our only excuse is that we shall make them explicit rather than hide them under the rug. And we shall argue that they make physical sense. The main tool will be *coarse-graining*.

An earlier version of this arduous road was shorter and not quite correct.[3]

2. THE CLASSICAL STARTING POINT

We first deal with a system of N classical particles in a volume Ω with reflecting walls. The $6N$ equations of motion in Hamiltonian form are

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}. \tag{1}$$

An ensemble of such systems is described by a density $\rho(q, p, t)$ in the $6N$-dimensional phase space. It obeys the Liouville equation

$$\dot{\rho}(q, p, t) = \{H, \rho\}, \tag{2}$$

where $\{\ \}$ are Poisson brackets. When ρ happens to depend on q and p only through the function $H(q, p)$ its time derivative (2) obviously vanishes. In particular a stationary ensemble with given energy E is

$$\rho(q, p) = \delta[H(q, p) - E]. \tag{3}$$

Incidentally, it is important to emphasize that the solution of the Liouville equation (2) is mathematically equivalent to the solution of (1) in the following sense. The solutions of (1) describe the path along which any point in phase space travels; if one knows all these paths one also knows how a swarm of points with density ρ will develop in time. Vice versa, if one knows the solution of (2) for arbitrary initial ρ it implies that one knows the motion if ρ is a δ-function at a single point. Introducing an ensemble is therefore not an aid to solving the equations of motion but merely a preliminary step for an approximate treatment to come.

3. COARSE-GRAINING

Let the set of slow variables which are subject to observation consist of the functions $A_r(q, p)$. Define *phase cells* as regions of phase space delineated by

$$a_r < A_r(q, p) < a_r + \Delta a_r \qquad \text{(all } r\text{)}. \tag{4}$$

The size of the margins Δa_r is not precisely fixed but at least they should be smaller than the inaccuracy of the observations. On the other hand, they must be so large that each cell contains a huge number of particles. These are indispensable strictures on the kind of system to which statistical mechanics applies, as was already clear from the works of Boltzmann[4] and Ehrenfest.[5] If the number of particles is not large enough the macroscopic equations are spoiled by the appearance of fluctuations.

The density ρ in phase space gives rise to a coarse-grained distribution P in the space of observables given by

$$P(a_1, a_2, ...) \, \Delta a_1 \, \Delta a_2 \cdots = \int_{a_r < A_r(q, p) < a_r + \Delta a_r} \rho(q, p) \, dq \, dp. \tag{5}$$

This is a projection in the sense that from ρ follows uniquely the density P but not vice versa. Any stationary ρ^e projects into a stationary P^e.

4. THE MASTER EQUATION

From the microscopic point of view the strict equation of motion (2) for ρ does not project into a differential equation of motion for P, owing to the elimination of the majority of variables. The distribution P as a function of time does not form a semi-group because the information about the details of the distribution ρ in the interior of each cell has been lost. This is remedied by the usual device of physicists: Make an assumption. One *assumes* that the details of the interior distribution do not matter; all that matters for the evolution of P is the total occupation in each cell as given by (5) and therefore it is justifiable to smear out in each cell the density ρ so as to make it constant in the cell. It is then possible to utilize (2) and compute ρ a short time $\varDelta t$ later. Take a cell

$$\varDelta a = \{a_1, a_1 + \varDelta a_1; a_2, a_2 + \varDelta a_2; ...\}; \tag{6}$$

the fraction of the interior of the cell that has fanned out into another cell $\varDelta a'$ we denote by $\varDelta t . W(a' \mid a) \, \varDelta a'$. It is then clear that the distribution over the various phase cells changes with time according to

$$\dot{P}(a; t) = -\int W(a' \mid a) \, da' \, \dot{P}(a; t) + \int W(a \mid a') \, P(a'; t) \, da'. \tag{7}$$

This is what has been called the "master equation"[6] The transition probability $W(a' \mid a)$ represents the probability per unit time for jumping from a to a'.

The first and second moments of the jumps are the mean and covariance,

$$m_r(a) = \int (a'_r - a_r) \, W(a' \mid a) \, da', \tag{8}$$

$$\sigma_{rs}(a) = \int (a'_r - a_r)(a'_s - a_s) \, W(a' \mid a) \, da'. \tag{9}$$

On formally expanding P in the master equation (7) one obtains the "Kramers–Moyal series."[7] The first two terms are

$$\dot{P}(a, t) = -\sum_r \frac{\partial}{\partial a_r} m_r(a) \, P(a, t) + \frac{1}{2} \sum_{r, s} \frac{\partial^2}{\partial a_r \, \partial a_s} \sigma_{rs}(a) \, P(a, t) + \cdots \tag{10}$$

However, this is *not* a systematic expansion, as we shall see.[8]

5. THE REPEATED RANDOMNESS ASSUMPTION

The same randomness assumption is needed after each Δt. This repeated randomness assumption is a generalization of Boltzmann's "Stosszahlansatz." It is part of *all* derivations of irreversible processes, even though it may be cleverly concealed.[9] It breaks the time symmetry by explicitly postulating the randomization *at the beginning* of the time interval Δt. There is no logical justification for this assumption other than that it is the only thing one can do and that it works. If one assumes randomness at the end of each Δt coefficients for diffusion, viscosity, etc. appear with the wrong sign; if one assumes randomness at the midpoint no irreversibility appears.

The proper choice of the macrovariables A_r is crucial; it is not determined by the taste of the experimenter. These variables must somehow incorporate all correlations that live longer than Δt. For instance, in the hydrodynamics of a mixture one must include the local composition among the A_r. On the other hand, any variable that does adjust itself rapidly to the local equilibrium need not be included as a macrovariable by itself, for its value is a function of the others. This is the reason why one cannot improve on thermodynamics by adding new thermodynamic variables, as is done in "extended thermodynamics."[10]

6. THE QUANTUM MECHANICAL BASIS

In quantum mechanics the basic equation is, instead of (1), the Schrödinger equation for all N particles of the system :

$$i\hbar\dot{\Psi}(q_1, q_2,...; t) = H\Psi(q_1, q_2,...; t). \tag{11}$$

Statistical mechanics deals with macroscopic systems having numerous particles. It should be realized that in such systems the density of levels is enormous. As a consequence the system is never in a single eigenstate but always in a superposition of a very large number of eigenstates. *The usual picture of an atom being in one or the other eigenstate is inappropriate for macroscopic systems*. This is the reason why the typical quantum mechanical features as coherence and entanglement do not show up in our macroscopic world. In particular, each energy shell consists of the large subspace spanned by the eigenfunctions of (11) belonging to all eigenvalues between E and $E + \Delta E$.

The macrovariables A_r are those hermitian operators in the Hilbert space of the Ψ that correspond to quantities that we observe as slowly varying. That implies that they are almost diagonal in the representation in

which H is diagonal. More precisely, in this representation their matrix elements are concentrated in a narrow strip along the diagonal; any matrix element that connects two eigenvalues of E with a difference of more than ΔE may be neglected. Hence the operator A_1 may be diagonalized in each shell separately, but it still depends on time. It is then possible to continue this process for A_2, A_3,..., etc. In this way one ends up with phase cells in which not only the energy but also the variables A_r have well-defined values a_r within the margins of experimental uncertainty. Nonetheless they are still many-dimensional subspaces of Hilbert space.[11]

   This construction is the quantum analog of coarse-graining. The probability for the system to be in the phase cell $P(a_1, a_2,...) \, \Delta a_1 \, \Delta a_2 \cdots$ is equal to the absolute square $|\mathbb{P}\Psi|^2$ of the projection $\mathbb{P}$ of Ψ into that cell. It can be argued, though not proved, that P again obeys the same master equation (7) under the same restrictions as in the classical case. Roughly speaking, the quantum character has endowed the phase space with a fine-grained structure determined by $\hbar$, but that does not affect the already existing coarse grains. Of course, the actual values of the transition probabilities W have to be computed by using the actual Schrödinger equation.

7. TIME SYMMETRY

   In the classical language the *time reversal transformation* reads

$$t \rightarrow -t, \quad q \rightarrow q, \quad p \rightarrow -p. \tag{12}$$

Suppose that under this transformation A_r is either even or odd:

$$A_r(q, -p) = \varepsilon_r A_r(q, p), \tag{13}$$

where $\varepsilon_r = \pm 1$ for the even and the odd case. Furthermore H must be even so that the equilibrium distribution is invariant: $P^e(a) = P^e(\varepsilon a)$. Note that this excludes the case of an external magnetic field, which could be included at the expense of a slightly more complicated formalism. It can be proved rigorously that the transition probabilities in the master equation (7) have the symmetry property

$$W(a \mid a') \, P^e(a') = W(\varepsilon a' \mid \varepsilon a) \, P^e(a), \tag{14}$$

which is called "detailed balance." This property is the basis for the reciprocity relations. In quantum mechanics a slightly different formulation yields the same result.[11]

8. THE EXPANSION OF THE MASTER EQUATION

So far we have performed the first major step, the step from microscopic equations for molecules to the description in terms of phase cells and the transition probabilities between them. This level of description was originally called *mesoscopic*, although that term has later been used for other purposes. We now have to extract from the mesoscopic master equation the familiar *macroscopic* deterministic equations that are used to describe phenomena in the laboratory such as hydrodynamics, Ohm's law, rate equations of chemical reactions, etc. This is achieved by the *system size expansion*.[8, 12] It applies whenever W involves a large parameter Ω, for instance the volume V of a gas, the capacity C of a condenser, or a particle number N,

$$W(a \mid a') = W^{(0)} + \Omega^{-1}W^{(1)} + \Omega^{-2}W^{(2)} + \cdots \tag{15}$$

(There may be an overall factor Ω, but we omit it as it may be absorbed into the unit of time.) Accordingly the mean and variance in (8), (9) may be written

$$m_r(a) = m_r^{(0)}(a) + \Omega^{-1}m_r^{(1)}(a) + \cdots \tag{16}$$

$$\sigma_{rs}(a) = \sigma_{rs}^{(0)}(a) + \cdots \tag{17}$$

Next, assuming that the a_r were chosen to be extensive variables, we may write them as the sum of a macroscopic part proportional to Ω and a fluctuating part of order $\Omega^{\frac{1}{2}}$

$$a_r = \Omega\phi_r(t) + \Omega^{\frac{1}{2}}\xi, \qquad P(a, t) = \Pi(\xi, t). \tag{18}$$

This is to be considered as a time dependent transformation from the variables a_r to the new variables ξ_r, in which certain functions $\phi_r(t)$ occur, which will be chosen presently. The mean and variance of ξ_r defined in (16), (17) become to lowest order

$$m_r(\phi) = m_r^{(0)}(\phi) + \Omega^{-\frac{1}{2}} \sum_s m_{r|s}^{(0)}(\phi)\, \xi_s + \Omega^{-1}m_r^{(1)}(\phi) + \cdots \tag{19}$$

$$\sigma_{rs}(\phi) = \sigma_{rs}^{(0)}(\phi) + \mathcal{O}(\Omega^{-\frac{1}{2}}). \tag{20}$$

Here m_r and σ_{rs} are regarded as functions of the intensive variables ϕ while $m_{r|s}$ is the derivative with respect to ϕ_s.

Applying this expansion to the master equation (7) one gets

$$\frac{\partial \Pi(\xi, t)}{\partial t} - \sum_r \frac{\partial \Pi(\xi, t)}{\partial \xi_r} \Omega^{\frac{1}{2}} \frac{d\phi_r}{dt}$$

$$= -\Omega^{\frac{1}{2}} \sum_r m_r^{(0)} \frac{\partial \Pi}{\partial \xi_r} - \sum_{r,s} m_{r|s}^{(0)} \frac{\partial}{\partial \xi_r} \xi_s \Pi + \frac{1}{2} \sum_{r,s} \sigma_{rs}^{(0)} \frac{\partial^2 \Pi}{\partial \xi_r \partial \xi_s}. \tag{21}$$

The two terms involving the large factor $\Omega^{\frac{1}{2}}$ can be made to cancel by choosing the functions $\phi_r(t)$ such as to obey

$$\frac{d\phi_r}{dt} = m_r^{(0)}(\phi). \tag{22}$$

These are the (nonlinear) macroscopic equations. The equilibrium ϕ_r^e is found by solving $m_r^{(0)}(\phi^e) = 0$. It has been implicitly supposed that the equations (22) are such that the solution $\phi_r(t)$ tends to ϕ^e as $t \to \infty$.

9. THE LINEAR REGIME

Near equilibrium one may put

$$\phi_r = \phi_r^e + \eta_r, \qquad \frac{d\eta_r}{dt} = \sum_s m_{r|s}^{(0)}(\phi^e)\, \eta_s. \tag{23}$$

These are the linear regression equations. In order to write them in the usual Onsager formulation, we first notice from (21) that the equilibrium distribution Π^e obeys in the same linear regime

$$-\sum_s m_{r|s}^{(0)} \eta_s + \frac{1}{2} \sum_s \sigma_{rs}^{(0)} \frac{\partial \log \Pi^e}{\partial \eta_s} = 0. \tag{24}$$

The Onsager affinities are defined by

$$X_s = \partial \log \Pi^e / \partial \eta_s. \tag{25}$$

Hence the linear regression equations (23) may be written

$$\frac{d\eta_r}{dt} = \frac{1}{2} \sum_s \sigma_{rs}^{(0)} X_s. \tag{26}$$

The Onsager coefficients are

$$L_{rs} = \tfrac{1}{2}\sigma(\phi^e);$$ (27)

they are symmetric as is manifest from (9).

Note. In the previous version[3] it was stated that the nonlinear equations (22) did not give rise to any reciprocity relations, but we now find that their linear approximation does. In fact, the present derivation closely parallels the original one by Onsager.

10. THE DIFFUSION TYPE (REF. 13)

In the preceding application of the Ω-expansion it was assumed that the solutions $\phi_r(t)$ of (22) tend to an equilibrium value ϕ^e. Obviously this is not true in the case that the $m_r^{(0)}$ are all identically equal to zero. In that case there is no macroscopic approach to equilibrium to keep the fluctuations in check. The expansion of the master equation (7) now takes the radically different form

$$\frac{\partial P(a, t)}{\partial t} = -\sum_r \frac{\partial}{\partial a_r} m_r^{(1)}(a)\, P + \frac{1}{2} \sum_{rs} \frac{\partial^2}{\partial a_r\, \partial a_s} \sigma_{rs}^{(0)}(a)\, P.$$ (28)

This is a Fokker–Planck equation for the distribution P of a type which has been called "nonlinear" as the coefficients are nonlinear functions of the independent variables a_r.

By a rearrangement of terms it may be written in a form more suitable for our purpose:

$$\frac{\partial P}{\partial t} = -\sum_r \frac{\partial}{\partial a_r} K_r(a)\, P + \frac{1}{2} \sum_{r,s} \frac{\partial}{\partial a_r} \sigma_{rs}^{(0)}(a)\, P^e \frac{\partial}{\partial a_s} \frac{P}{P^e}$$ (29)

$$K_r(a) = m_r^{(1)}(a) - \frac{1}{2P^e} \sum_s \frac{\partial}{\partial a_s} \sigma_{rs}^{(0)}(a)\, P^e.$$ (30)

The first term of (29) represents a Liouville equation belonging to the deterministic equation of motion

$$\dot{a}_r = K_r(a).$$ (31)

The second term is purely dissipative, as is demonstrated by the following H-theorem

$$\frac{d}{dt} \int \frac{P^2}{P^e}\, da = -\sum_{r,s} \int P^e \left(\frac{\partial}{\partial a_r} \frac{P}{P^e}\right) \sigma_{rs}^{(0)}(a) \left(\frac{\partial}{\partial a_s} \frac{P}{P^e}\right) da \leqslant 0.$$ (32)

The integral kernel $W(a \,|\, a')$ in (7) has now been replaced with the differential operator occurring in (29). In order to apply to it the detailed balance relation (14) we reformulate this relation by means of two test functions $f(a)$, $g(a)$:

$$\int da \int da' \, f(a) \, \frac{W(a \,|\, a')}{P^e(a)} \, g(a') = \int da \int da' \, f(a) \, \frac{W(\varepsilon a' \,|\, \varepsilon a)}{P^e(a')} \, g(a').$$

Now insert (29) and perform the necessary partial integrations; the result is after some manipulations[12]

$$\varepsilon_r K_r(\varepsilon a) = -K_r(a), \qquad \varepsilon_r \varepsilon_s \sigma_{rs}^{(0)}(\varepsilon a) = \sigma_{rs}^{(0)}(a). \tag{33}$$

One sees that the two terms in (29) are symmetric and antisymmetric in time respectively.

11. THE SECOND EXPANSION

In order to extract a macroscopic feature one has to scale down the fluctuations represented by the second term in (29). For this purpose it is necessary to pick a new expansion parameter. An obvious choice is the temperature since one expects that the fluctuations disappear at low temperature. Accordingly we assume that $\sigma_{rs}^{(0)}$ involves a factor T:

$$\sigma_{rs}^{(0)} = T \tilde{\sigma}_{rs}. \tag{34}$$

For any particular system this can be checked using the definition of σ_{rs}. We also know how P^e varies with T:

$$P^e(a) = \text{const. } \exp[-F(a)/T], \tag{35}$$

where F is the free energy. On substituting (34) and (35) into (29) and collecting the powers of T one gets

$$\frac{\partial P}{\partial t} = -\sum_r \frac{\partial}{\partial a_r} \left[K_r - \frac{1}{2} \sum_s \tilde{\sigma}_{rs} \frac{\partial F}{\partial a_s} \right] P + \frac{1}{2} T \frac{\partial}{\partial a_r} \tilde{\sigma}_{rs} \frac{\partial P}{\partial a_s}. \tag{36}$$

12. THE NONLINEAR ONSAGER RELATIONS

Evidently in the limit $T \to 0$ the second part vanishes and one is left with an equation of the form of a Liouville equation corresponding with a deterministic equation for a_r:

$$\dot{a}_r = K_r(a) - \frac{1}{2} \sum_s \tilde{\sigma}_{rs}(a) \frac{\partial F}{\partial a_s}. \tag{37}$$

In this equation the first term is reversible according to (33) and therefore of mechanical nature. The second term is the damping. The derivatives of the free energy are Onsager's affinities X_s. In contrast to the standard case they need not be linear combinations of the a_r. The usual coefficients L_{rs} are here replaced with $-\frac{1}{2}\tilde{\sigma}_{rs}(a)$; they also may be nonlinear functions of the a_r. Yet they are symmetric; these are the nonlinear Onsager relations. We have derived them for systems of diffusion type. For details see ref. 12, Chapter XI.

REFERENCES

1. L. Onsager, *Phys. Rev.* **37**:405 and **38**:2265 (1931). H. B. G. Casimir, *Rev. Mod. Phys.* **17**:343 (1945). S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962).
2. R. Benguria and M. Kac, *Phys. Rev.* **46**:1 (1981). H. Grabert, U. Weiss, and P. Talkner, *Z. Phys. B* **55**:87 (1984). P. Mazur showed in *Periodic Polytechnica, Ser. Chem. Eng.* **41** (2) (1997) that Onsager came close.
3. N. G. van Kampen, *Periodica Polytechnica, Ser. Chem. Eng.* **41**(2):157 (1997).
4. L. Boltzmann, *Vorlesungen über Gastheorie* (Barth, Leipzig, 1896 & 1898).
5. P. and T. Ehrenfest, *Begriffliche Grundlagen der Statistischen Auffassung in der Mechanik*, Encyklopädie der Mathematischen Wissenschaften IV, Vol. 32 (Teubner, Leipzig, 1912).
6. A. Nordsieck, W. E. Lamb, and G. E. Uhlenbeck, *Physica* **7**:344 (1940).
7. H. A. Kramers, *Physica* **7**:284 (1940). J. E. Moyal, *J. Roy. Statist. Soc. B* **11**:150 (1949).
8. N. G. van Kampen, *Can. J. Phys.* **39**:551 (1961).
9. L. van Hove, *Physica* **21**:517 (1955); **23**:441 (1957).
10. I. Müller, *Z. Phys.* **198**:329 (1967). R. E. Nettleton, *Phys. Fluids* **2**:256 (1959); **3**:216 (1960).
11. N. G. van Kampen, *Physica* **20**:603 (1954).
12. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 2nd Ed. (North-Holland, 1992), Section XVII, **7**.
13. N. G. van Kampen, *Phys. Lett. A* **62**:383 (1977).