

# Elementary Derivation of Nonlinear Transport Equations from Statistical Mechanics

John D. Ramshaw<sup>1</sup>

*Received April 30, 1986; revision received August 5, 1986*

---

Exact closed nonlinear transport equations for a set of macroscopic variables  $\mathbf{a}$  are derived from classical statistical mechanics. The derivation involves only simple manipulations of the Liouville equation, and makes no use of projection operators or graphical expansions. It is based on the Chapman–Enskog idea of separating the distribution function into a constrained equilibrium part, obtained from information theory, and a small remainder. The resulting exact transport equations involve time convolutions over the past history of both  $\mathbf{a}(t)$  and  $\dot{\mathbf{a}}(t)$ . However, if the variables  $\mathbf{a}$  provide a complete macroscopic description, the equations may be simplified. This is accomplished by a systematic expansion procedure of Chapman–Enskog type, in which the small parameter is the natural parameter of slowness relevant to the problem. When carried out to second order, this expansion leads to approximate nonlinear transport equations that are local in time. These equations are valid far from equilibrium. They contain nonlinear (i.e., state-dependent) transport coefficients given by integrals of time correlation functions in the constrained equilibrium ensemble. Earlier results are recovered when the equations are linearized about equilibrium. As an illustrative application of the formalism, an expression is derived for the nonlinear (i.e., velocity-dependent) friction coefficient for a heavy particle in a bath of light particles.

---

**KEY WORDS:** Nonlinear transport equations; nonlinear evolution equations; transport far from equilibrium; transport coefficients; information theory; Chapman–Enskog theory; closure.

## 1. INTRODUCTION AND SUMMARY

A primary objective of nonequilibrium statistical mechanics is the derivation from microscopic first principles of macroscopic transport

---

<sup>1</sup>Theoretical Division, University of California, Los Alamos National Laboratory, Los Alamos, New Mexico 87545. Present address: Advanced Methods Projects, EG & G Idaho, Inc., Idaho National Engineering Laboratory, P.O. Box 1625, Idaho Falls, Idaho 83415.

equations; i.e., closed equations of motion for reduced sets of macroscopic variables. In general such equations are both nonlinear and irreversible (e.g., the Navier–Stokes equations of hydrodynamics). In much of the early work, attention was focused on the irreversible aspects within the simplified context of linearized transport equations. This work was largely concerned with deriving microscopic expressions for the linear transport coefficients.<sup>(1)</sup> In recent years the emphasis has shifted toward nonlinear descriptions of irreversible processes in systems far from equilibrium. This work has been pursued by a variety of different methods.<sup>(2–4)</sup>

A class of methods that is especially well suited to the derivation of closed deterministic transport equations is that based on the application of Chapman–Enskog ideas<sup>(5)</sup> to the Liouville equation.<sup>(6–10)</sup> In these methods the full distribution function is typically represented as the sum of a constrained or local equilibrium part and a presumably small remainder. (For general use we prefer “constrained” to “local,” as the latter is strictly appropriate only for spatially dependent macroscopic variables.) Perhaps the most elegant version of this approach is that of Robertson,<sup>(8)</sup> who obtains exact closed equations of motion with a very appealing structure.

Robertson’s theory makes use of time-dependent projection operators, which have both advantages<sup>(8)</sup> and disadvantages.<sup>(9)</sup> The projection operators confer certain desirable properties upon the correlation functions that arise.<sup>(8)</sup> The price paid is that the correlation functions do not directly reflect the true microscopic dynamics, but rather evolve according to an artificial “projected” dynamics.

Here we pursue an analogous development that uses no projection operators. Attention is restricted to the classical case for simplicity; the corresponding quantum development is straightforward.<sup>(8)</sup> Like Robertson, we use information theory to define the constrained equilibrium distribution, and we take the initial distribution function to be the constrained equilibrium distribution consistent with the initial values of the macroscopic variables. Simple manipulations of the Liouville equation then lead, in a remarkably straightforward way, to exact nonlinear transport equations for the macroscopic variables  $\mathbf{a}$ . Many of the earlier theories in the Chapman–Enskog spirit<sup>(6)</sup> are essentially based on a naive low-order iterative approximation to these exact equations.

The resulting exact transport equations, like those of Robertson, are nonlocal in time; they involve convolutions over the past history of the system. But while Robertson’s equations involve only past values of  $\mathbf{a}(t)$ , the present equations contain an additional convolution over past values of  $\dot{\mathbf{a}}(t)$  as well. This is a direct consequence of omitting the projection operator. Thus, the present equations are not merely integrodifferential in nature, but are also implicit in  $\dot{\mathbf{a}}(t)$ , whereas Robertson’s equations give

$\dot{\mathbf{a}}(t)$  explicitly in terms of the past history of  $\mathbf{a}(t)$ . This implicitness may appear to be a significant complication, but in practice it presents no real problems. The reason is that the exact transport equations are not normally solved in any case; they are primarily of interest as a structure upon which to impose approximations. As will be seen, the present equations are well suited for this purpose.

If the variables  $\mathbf{a}$  provide a complete macroscopic description, which is our hope and intent, then the exact transport equations must reduce to equations that are essentially local in time. They can then be solved as an initial-value problem, which is just what we require of a macroscopic state description. Physically, this reduction to time-local behavior is a consequence of a wide separation between the fast microscopic time scales and the slow time scales over which the variables  $\mathbf{a}$  change appreciably. What is needed is a way to express this presumed separation of time scales as a well-defined mathematical approximation in which the exact transport equations reduce to a time-local form.

Robertson's formulation has the advantage that the reduction to time-local equations is very easy to perform. One simply assumes that a certain kernel decays essentially to zero on a microscopic time scale, so that it may be macroscopically approximated as a delta function. However, the resulting transport coefficients still involve the projected dynamics, and if it is desired to relate them to the true microscopic dynamics, an additional conversion is required.<sup>(10,11)</sup>

In the present formulation the reduction to time-local equations is not quite so straightforward, as it occurs through an interplay between two convolution kernels, neither of which becomes nearly local by itself. Nevertheless, the reduction is readily accomplished by means of a systematic expansion procedure of Chapman–Enskog type,<sup>(5,7)</sup> in which the small parameter is the natural parameter of slowness<sup>(12)</sup> (call it  $\xi$ ) relevant to the problem. The limit  $\xi \rightarrow 0$  corresponds to an infinite separation of time scales; i.e., a macroscopic time scale infinitely slower than the microscopic one.

It is necessary to carry the expansion to second order to obtain irreversibility, as the first-order equations are purely reversible in nature. When this is done, the two nonlocal kernels combine in the term of order  $\xi^2$  to produce a correlation function containing a “subtracted flux.”<sup>(8)</sup> The transport equations then become local in time under the assumption that this correlation function decays essentially to zero on a microscopic time scale, so that it may be approximated as a delta function. (Of course, this presupposes that any long-time tails are integrable and small in amplitude; otherwise, a time-local representation of the irreversible terms does not strictly exist.) We thereby obtain nonlinear irreversible transport equations,

local in time and valid to order  $\xi^2$ , in which the transport coefficients are given in terms of time correlation functions involving the true microscopic dynamics in the constrained equilibrium ensemble. When these results are linearized about equilibrium, they reduce to results obtained in earlier analyses of linear transport.<sup>(12)</sup>

In summary, our development combines the following basic ingredients: (1) a Chapman–Enskog separation of the distribution function into a constrained equilibrium distribution and a remainder,<sup>(6, 10)</sup> (2) the use of information theory<sup>(13)</sup> to define the constrained equilibrium distribution,<sup>(8)</sup> and (3) a systematic expansion of Chapman–Enskog type<sup>(7)</sup> in the appropriate parameter of slowness.<sup>(12)</sup> The earlier work to which we are closest in spirit is that of Robertson,<sup>(8)</sup> Zwanzig,<sup>(12)</sup> Dufty,<sup>(7)</sup> Wong *et al.*,<sup>(14)</sup> and Dufty and Lindenfeld.<sup>(15)</sup> Our formulation is perhaps best thought of as a variation on Robertson’s theory, in which the projection operator is avoided by exploiting the presence in the problem of the small parameter  $\xi$ .

## 2. PRELIMINARIES

The microscopic state or phase of the system is denoted by  $\mathbf{X}$ , which represents the set of all canonical coordinates and momenta of the constituent particles. A set of macroscopic variables  $\mathbf{a} = (a_1, a_2, \dots)$  has somehow been identified which is believed or hoped to provide a complete description of the macroscopic state of the system. The corresponding microscopic dynamical variables are denoted by  $\mathbf{A}(\mathbf{X}) = (A_1(\mathbf{X}), A_2(\mathbf{X}), \dots)$ .

The time-dependent distribution function in phase space is denoted by  $\rho(\mathbf{X}, t)$ . It satisfies the Liouville equation

$$\partial\rho/\partial t = -L\rho \quad (1)$$

where  $L$  is the Liouville operator; i.e., the Poisson bracket of the operand with the Hamiltonian. This equation determines  $\rho(\mathbf{X}, t)$  in terms of its initial value  $\rho(\mathbf{X}, 0)$ . The macroscopic variables  $\mathbf{a}(t)$  at time  $t$  are then given by

$$\mathbf{a}(t) = \int d\mathbf{X} \rho(\mathbf{X}, t) \mathbf{A}(\mathbf{X}) \quad (2)$$

The dynamical variable corresponding to the time derivative  $\dot{\mathbf{a}}(t)$  is  $\dot{\mathbf{A}}(\mathbf{X}) = L\mathbf{A}(\mathbf{X})$ , so that

$$\dot{\mathbf{a}}(t) = \int d\mathbf{X} \rho(\mathbf{X}, t) \dot{\mathbf{A}}(\mathbf{X}) \quad (3)$$

Our next task is to define the constrained equilibrium distribution, which is intended to be an approximation to the true distribution that reproduces the same values of  $\mathbf{a}$ . It is natural to base this definition on information theory, which provides a distribution containing minimal additional information beyond the specified values of  $\mathbf{a}$ . This distribution is of the generalized canonical form<sup>(8,13)</sup>

$$\rho_0(\mathbf{X}; \mathbf{a}) = Q^{-1}(\mathbf{a}) \exp\{\gamma(\mathbf{a}) \cdot \mathbf{A}(\mathbf{X})\} \tag{4}$$

where

$$Q(\mathbf{a}) = \int d\mathbf{X} \exp\{\gamma(\mathbf{a}) \cdot \mathbf{A}(\mathbf{X})\} \tag{5}$$

and the Lagrange multipliers  $\gamma(\mathbf{a})$  are functions of  $\mathbf{a}$  implicitly determined by the requirement that

$$\int d\mathbf{X} \rho_0(\mathbf{X}; \mathbf{a}) \mathbf{A}(\mathbf{X}) = \mathbf{a} \tag{6}$$

If  $Q$  is alternatively regarded as a function of  $\gamma$ , one readily verifies that

$$\mathbf{a} = \frac{\partial}{\partial \gamma} \ln Q(\gamma) \tag{7}$$

This determines a functional relation  $\mathbf{a}(\gamma)$ , which may in principle be inverted to obtain  $\gamma(\mathbf{a})$ . The deviations of the parameters  $\gamma$  from their equilibrium values play the role of thermodynamic driving forces in the macroscopic dynamics.<sup>(8)</sup>

Averages weighted by  $\rho_0(\mathbf{X}; \mathbf{a})$  will be denoted by angular brackets; i.e.,  $\langle F(\mathbf{X}) \rangle = \int d\mathbf{X} \rho_0(\mathbf{X}; \mathbf{a}) F(\mathbf{X})$ . It must be remembered that such averages depend parametrically upon  $\mathbf{a}$ , as this is not explicitly indicated by the bracket notation. When it is necessary to specify the time at which  $\mathbf{a}$  is evaluated, a subscript will be affixed; e.g.,  $\langle F(\mathbf{X}) \rangle_t = \int d\mathbf{X} \rho_0(\mathbf{X}; \mathbf{a}(t)) F(\mathbf{X})$ . We emphasize that the angular brackets always denote a constrained equilibrium average and never refer to an average weighted by the full distribution function  $\rho(\mathbf{X}, t)$ . Averages of the latter type will always be written out explicitly, as in Eqs. (2) and (3).

The constrained equilibrium approximation to  $\dot{\mathbf{a}}$  is just  $\dot{\mathbf{a}} \approx \mathbf{v}(\mathbf{a})$ , where

$$\mathbf{v}(\mathbf{a}) = \int d\mathbf{X} \rho_0(\mathbf{X}; \mathbf{a}) \dot{\mathbf{A}}(\mathbf{X}) = \langle \dot{\mathbf{A}}(\mathbf{X}) \rangle \tag{8}$$

is the canonical streaming velocity. It does not involve the microscopic dynamics, so its evaluation is simply a problem in equilibrium statistical

mechanics. In this approximation the macroscopic dynamics is purely reversible.<sup>(8)</sup>

Standard Liouville operator manipulations<sup>(10,11)</sup> enable us to rewrite  $\mathbf{v}(\mathbf{a})$  as

$$\mathbf{v}(\mathbf{a}) = -\langle \mathbf{A}\dot{\mathbf{A}} \rangle \cdot \boldsymbol{\gamma}(\mathbf{a}) \quad (9)$$

In a similar way, one readily verifies that

$$\mathbf{v}(\mathbf{a}) \cdot \boldsymbol{\gamma}(\mathbf{a}) = 0 \quad (10)$$

Equation (9) can therefore be written in the alternative form

$$\mathbf{v}(\mathbf{a}) = -\langle \delta \mathbf{A}\dot{\mathbf{A}} \rangle \cdot \boldsymbol{\gamma}(\mathbf{a}) \quad (11)$$

where  $\delta \mathbf{A}(\mathbf{X}) = \mathbf{A}(\mathbf{X}) - \mathbf{a}$ . This form will be useful later.

The entropy of the macroscopic state  $\mathbf{a}$  is given by<sup>(8)</sup>

$$S(\mathbf{a}) = -k_B \int d\mathbf{X} \rho_0(\mathbf{X}; \mathbf{a}) \ln \rho_0(\mathbf{X}; \mathbf{a}) = k_B [\ln Q(\mathbf{a}) - \mathbf{a} \cdot \boldsymbol{\gamma}(\mathbf{a})] \quad (12)$$

where  $k_B$  is Boltzmann's constant. It follows that

$$\partial S(\mathbf{a}) / \partial \mathbf{a} = -k_B \boldsymbol{\gamma}(\mathbf{a}) \quad (13)$$

which shows that  $\boldsymbol{\gamma}$  is thermodynamically conjugate to  $\mathbf{a}$ . It further follows that

$$\dot{S}(\mathbf{a}) = -k_B \boldsymbol{\gamma}(\mathbf{a}) \cdot \dot{\mathbf{a}} \quad (14)$$

The reversibility of  $\mathbf{v}(\mathbf{a})$  may be inferred directly from Eqs. (10) and (14), which show that  $\mathbf{v}(\mathbf{a})$  makes no contribution to the production of entropy.

### 3. EXACT NONLINEAR TRANSPORT EQUATIONS

We begin the derivation by writing

$$\rho(\mathbf{X}, t) = \rho_0(\mathbf{X}; \mathbf{a}(t)) + \delta\rho(\mathbf{X}, t) \quad (15)$$

which defines  $\delta\rho$ . Equation (3) then becomes

$$\dot{\mathbf{a}}(t) = \mathbf{v}(\mathbf{a}(t)) + \int d\mathbf{X} \delta\rho(\mathbf{X}, t) \dot{\mathbf{A}}(\mathbf{X}) \quad (16)$$

in which all irreversible effects reside in the term involving  $\delta\rho$ .

To determine  $\delta\rho$ , we combine Eqs. (1) and (15) to obtain

$$\left(\frac{\partial}{\partial t} + L\right) \delta\rho(\mathbf{X}, t) = -\left(\frac{\partial}{\partial t} + L\right) \rho_0(\mathbf{X}; \mathbf{a}(t)) \quad (17)$$

This equation may be solved by standard methods, with the result

$$\delta\rho(\mathbf{X}, t) = e^{-Lt} \left[ \delta\rho(\mathbf{X}, 0) - \int_0^t d\tau e^{L\tau} \left(\frac{\partial}{\partial \tau} + L\right) \rho_0(\mathbf{X}; \mathbf{a}(\tau)) \right] \quad (18)$$

We now make the usual assumption<sup>(8, 10)</sup> that the initial distribution is just the constrained equilibrium distribution consistent with the initial values of  $\mathbf{a}$ , so that  $\delta\rho(\mathbf{X}, 0)$  vanishes. This assumption has sometimes been a cause for concern, but it is quite easy to justify from the standpoint of information theory. If our knowledge of the initial conditions is in fact limited to the initial values of  $\mathbf{a}$ , then an initial constrained equilibrium distribution is the least biased or maximally noncommittal probability distribution consistent with that knowledge.<sup>(13)</sup> Moreover, if the variables  $\mathbf{a}$  do in fact constitute a complete macroscopic description, then their initial values are all that is needed to predict their later values. This in turn implies that the initial constrained equilibrium distribution contains all the information necessary to make such predictions,<sup>(13)</sup> and hence to derive the macroscopic equations of motion.

With the above choice of the initial distribution, Eq. (18) reduces to

$$\delta\rho(\mathbf{X}, t) = -\int_0^t d\tau e^{L(\tau-t)} \left(\frac{\partial}{\partial \tau} + L\right) \rho_0(\mathbf{X}; \mathbf{a}(\tau)) \quad (19)$$

The next step is to carry out the indicated operations in the integrand. We readily find that

$$\frac{\partial}{\partial t} \rho_0(\mathbf{X}; \mathbf{a}(t)) = \rho_0(\mathbf{X}; \mathbf{a}(t)) [\mathbf{A}(\mathbf{X}) - \mathbf{a}(t)] \cdot \mathbf{M}(\mathbf{a}(t)) \cdot \dot{\mathbf{a}}(t) \quad (20)$$

where

$$\mathbf{M}(\mathbf{a}) = \partial\gamma(\mathbf{a})/\partial\mathbf{a} = \langle \delta\mathbf{A} \delta\mathbf{A} \rangle^{-1} \quad (21)$$

and

$$L\rho_0(\mathbf{X}; \mathbf{a}) = \rho_0(\mathbf{X}; \mathbf{a}) \gamma(\mathbf{a}) \cdot \dot{\mathbf{A}}(\mathbf{X}) \quad (22)$$

Combining Eqs. (19)–(22) and substituting the resulting expression for  $\delta\rho$  into Eq. (16), we finally obtain

$$\dot{\mathbf{a}}(t) = \mathbf{v}(\mathbf{a}(t)) - \int_0^t d\tau [\mathbf{K}(t, \tau) \cdot \gamma(\mathbf{a}(\tau)) + \mathbf{L}(t, \tau) \cdot \dot{\mathbf{a}}(\tau)] \quad (23)$$

where

$$\mathbf{K}(t, \tau) = \langle [e^{L(t-\tau)} \dot{\mathbf{A}}(\mathbf{X})] \dot{\mathbf{A}}(\mathbf{X}) \rangle_{\tau} \quad (24)$$

$$\mathbf{L}(t, \tau) = \langle [e^{L(t-\tau)} \dot{\mathbf{A}}(\mathbf{X})] \delta \mathbf{A}(\mathbf{X}) \rangle_{\tau} \cdot \mathbf{M}(\mathbf{a}(\tau)) \quad (25)$$

and use has been made of the fact that  $e^{Lt}$  is an orthogonal operator.

The exact nonlinear transport equations are given by Eq. (23). They are integrodifferential equations which relate  $\mathbf{a}(\tau)$  and  $\dot{\mathbf{a}}(\tau)$  over the time interval  $0 \leq \tau \leq t$ . The correlation functions in the kernels  $\mathbf{K}$  and  $\mathbf{L}$  are evaluated in the constrained equilibrium ensemble. Thus, they depend parametrically on  $\mathbf{a}$ , which is to be evaluated as  $\mathbf{a}(\tau)$  as indicated. Their time dependence is generated by the true microscopic dynamics via the time-development operator  $e^{Lt}$ . Equations (23)–(25) are the foundation upon which the remainder of our development is based. Similar equations have been derived by Wong *et al.*<sup>(14)</sup> and Dufty and Lindenfeld<sup>(15)</sup> for the case in which the variables  $\mathbf{A}$  are continuous field variables representing densities of conserved quantities.

Constrained-equilibrium time correlation functions of the type in Eqs. (24) and (25) must be handled with care, as some of the familiar rules for manipulating equilibrium time correlation functions no longer apply. In particular, the constrained equilibrium state is not stationary (i.e.,  $L\rho_0 \neq 0$ ), so the time origin cannot be shifted at will. The correlation functions therefore depend separately on  $t$  and  $\tau$ , and not merely on  $t - \tau$ . Since the time origin is no longer arbitrary, it is no longer appropriate to think of  $F(\mathbf{X})$  as  $F(t=0)$  and  $e^{Lt}F(\mathbf{X})$  as  $F(t)$  in the usual way. A little thought shows that the time to be associated with  $F(\mathbf{X})$  is the time at which the associated distribution over  $\mathbf{X}$  is evaluated. Thus, in Eqs. (24) and (25) we may think of  $\dot{\mathbf{A}}(\mathbf{X})$  and  $\delta \mathbf{A}(\mathbf{X})$  as  $\dot{\mathbf{A}}(\tau)$  and  $\delta \mathbf{A}(\tau)$ , respectively. Now the operator  $e^{Lt}$  still *shifts* the time by  $t$ , in the sense that  $e^{Lt}F(\mathbf{X})$  is the value of  $F$  at time  $t + \tau$  in a system whose phase point was  $\mathbf{X}$  at time  $\tau$ . Thus, if  $F(\mathbf{X})$  is thought of as  $F(\tau)$ , then  $e^{Lt}F(\mathbf{X})$  is to be thought of as  $F(t + \tau)$ . Equations (24) and (25) may therefore be rewritten in the more compact forms

$$\mathbf{K}(t, \tau) = \langle \dot{\mathbf{A}}(t) \dot{\mathbf{A}}(\tau) \rangle_{\tau} \quad (26)$$

$$\mathbf{L}(t, \tau) = \langle \dot{\mathbf{A}}(t) \delta \mathbf{A}(\tau) \rangle_{\tau} \cdot \mathbf{M}(\mathbf{a}(\tau)) \quad (27)$$

which also makes the significance of the correlation functions clearer.



#### 4. REDUCTION TO TIME-LOCAL EQUATIONS

The exact transport equations of the previous section are primarily of interest as a starting point for approximations. As discussed in Section 1, we wish to formulate a proper mathematical approximation corresponding to the physical separation of time scales that gives rise to time-local macroscopic behavior.

The structure of Eq. (23) suggests that successive approximations to  $\dot{\mathbf{a}}(t)$  be generated simply by iteration, using  $\mathbf{v}(\mathbf{a})$  as a first approximation. [The same sequence of approximations is generated by a naive Chapman–Enskog expansion of  $\dot{\mathbf{a}}(t)$  in powers of a “small” parameter introduced into Eq. (23) as a factor in front of the integral.] The second approximation to  $\dot{\mathbf{a}}(t)$  would then be given by Eq. (23) with  $\dot{\mathbf{a}}(\tau)$  replaced by  $\mathbf{v}(\mathbf{a}(\tau))$  in the right member. It is just this approximation (or the linearization thereof) upon which the older transport theories of Chapman–Enskog type were based.<sup>(6)</sup>

Unfortunately, this simple and intuitive approximation procedure is now known to be unsatisfactory. The problem is that the macroscopic and microscopic dynamics are not treated consistently at each level of approximation, and as a result the reduction to time-local behavior does not properly occur.<sup>(7)</sup> The need for consistency arises from the fact that the time correlation functions contain slowly varying parts corresponding to the slow variation of the variables  $\mathbf{a}(t)$ .<sup>(7,12)</sup> Any approximation to the macroscopic equations governing  $\mathbf{a}(t)$  must therefore be accompanied by a self-consistent approximation in the slowly varying parts of the correlation functions.

The required consistency may be obtained by performing a more systematic expansion of Chapman–Enskog type in the natural parameter of slowness<sup>(12)</sup>  $\xi$  relevant to the problem. The existence of such a parameter is the usual reason for the separation of time scales and the slow variation of  $\mathbf{a}(t)$ . Typically  $\dot{\mathbf{a}} \sim \xi$  for small  $\xi$ ,<sup>(10)</sup> so that  $\dot{\mathbf{a}}(t)$  is small and  $\mathbf{a}(t)$  varies slowly.

The physical significance of  $\xi$  will be different in different contexts,<sup>(12)</sup> depending upon the physical origin of the separation of time scales. The most common origins are microscopic conservation laws and continuous broken symmetries.<sup>(11)</sup> Slow behavior can also result from large inertia or weak coupling.<sup>(12)</sup> Even when the physical origin of the slowness is apparent, the appropriate parameter of slowness may not be obvious on inspection. It may then be necessary to introduce  $\xi$  artificially, and subsequently restore it to unity after the expansion has been carried out. In such cases it is essential that  $\xi$  be introduced directly into the microscopic dynamics. It will then find its way into Eq. (23) via the microscopic

expressions for  $\mathbf{v}$ ,  $\mathbf{K}$ , and  $\mathbf{L}$ , and the required self-consistency will be automatic.

Occasionally one may encounter a problem with more than one parameter of slowness, and the present procedure will then need to be generalized. However, the restriction to a single parameter of slowness seems to be adequate in most cases. It should be noted that this restriction does *not* imply that there is only one slow time scale in the problem. The macroscopic evolution of  $\mathbf{a}(t)$  can still occur on several slow time scales  $\xi t$ ,  $\xi^2 t$ , etc. For example, in linearized hydrodynamics  $\xi$  may be identified with the wavenumber  $k$ , and the ideal (Euler) and irreversible (viscous) terms evolve on the time scales  $kt$  and  $k^2 t$ , respectively.

We therefore proceed to develop a systematic Chapman–Enskog expansion of  $\dot{\mathbf{a}}(t)$  in powers of  $\xi$ , based on Eq. (23) and the assumption that  $\dot{\mathbf{A}} \sim \xi$  for small  $\xi$ . In any particular application, the parameter  $\xi$  must be chosen so that this assumption is satisfied. (Notice that according to this assumption, our  $\xi$  corresponds to the square root of Zwanzig’s  $\xi$ .<sup>(12)</sup>) During the expansion the  $\mathbf{a}(t)$  are still to be regarded as having their correct values, as given by Eq. (2). Thus, they are formally regarded as quantities of order unity and are not expanded in  $\xi$ .<sup>(5)</sup> Their thermodynamic conjugates  $\gamma(\mathbf{a})$  are treated in the same way.

The leading dependence of the quantities  $\mathbf{v}$ ,  $\mathbf{K}$ , and  $\mathbf{L}$  upon  $\xi$  may be inferred from the number of factors of  $\dot{\mathbf{A}}$  that each contains. It follows that  $\mathbf{v} \sim \xi$ ,  $\mathbf{K} \sim \xi^2$ , and  $\mathbf{L} \sim \xi$  for small  $\xi$ . Inspection of Eq. (23) then shows that  $\dot{\mathbf{a}}(t)$  vanishes as  $\xi \rightarrow 0$ , as expected, so its expansion in  $\xi$  begins with the linear term. We therefore write

$$\dot{\mathbf{a}}(t) = \xi \dot{\mathbf{a}}_1(t) + \xi^2 \dot{\mathbf{a}}_2(t) + \dots \tag{28}$$

Substituting Eq. (28) into Eq. (23) and collecting coefficients, we obtain to order  $\xi^2$

$$\dot{\mathbf{a}}_1(t) = \lim_{\xi \rightarrow 0} [\xi^{-1} \mathbf{v}(\mathbf{a}(t))] \tag{29}$$

$$\dot{\mathbf{a}}_2(t) = \mathbf{v}_2(\mathbf{a}(t)) - \int_0^t d\tau \lim_{\xi \rightarrow 0} [\xi^{-2} \langle \dot{\mathbf{A}}(t) \Delta \dot{\mathbf{A}}(\tau) \rangle_\tau] \cdot \gamma(\mathbf{a}(\tau)) \tag{30}$$

where

$$\Delta \dot{\mathbf{A}}(\mathbf{X}) = \dot{\mathbf{A}}(\mathbf{X}) - \delta \mathbf{A}(\mathbf{X}) \cdot \mathbf{M}(\mathbf{a}) \cdot \langle \delta \mathbf{A} \dot{\mathbf{A}} \rangle \tag{31}$$

$\mathbf{v}_2(\mathbf{a})$  is the coefficient of  $\xi^2$  in the expansion of  $\mathbf{v}(\mathbf{a})$ , and use has been made of Eqs. (11), (26), and (27). Notice the automatic appearance of the “subtracted flux.”<sup>(8)</sup>

Since  $\mathbf{v}(\mathbf{a})$  is reversible in nature, the macroscopic equations of motion to order  $\xi$  are both strictly local in time and purely reversible. It is therefore necessary to carry the expansion to order  $\xi^2$  in order to obtain the lowest nonvanishing contributions of both reversible and irreversible character. The quantity  $\mathbf{v}_2$  in Eq. (30) represents a reversible contribution of order  $\xi^2$  to  $\dot{\mathbf{a}}(t)$ . This contribution, however, should ordinarily be negligible compared to the reversible contribution of order  $\xi$ , and it will therefore be neglected in what follows.

The reduction to time-local behavior in the irreversible term of order  $\xi^2$  is accomplished simply by assuming that the correlation function in Eq. (30) decays essentially to zero if  $\tau$  differs from  $t$  by more than some microscopic correlation time. This is a nontrivial assumption whose justification from first principles would require a detailed microscopic analysis of the correlation function. It will not always be satisfied, as counterexamples are known; e.g., the breakdown of hydrodynamics in two dimensions. However, the assumed behavior would seem to be necessary for the existence of a time-local macroscopic description of irreversibility. Experimental verification that such a description exists for a particular set of variables  $\mathbf{a}$  therefore constitutes an empirical justification for the assumption.

The above assumption, together with the fact that  $\mathbf{a}(t)$  is slowly varying, implies that the factor  $\gamma(\mathbf{a}(\tau))$  in Eq. (30) can be evaluated at  $\tau = t$  and taken outside the integral, and that all other  $\mathbf{a}(\tau)$ 's implicit in the correlation function can be replaced by  $\mathbf{a}(t)$  with negligible effect. Moreover, the lower integration limit can be extended to negative infinity. When all this is done and the integration variable is changed from  $\tau$  to  $s = t - \tau$ , Eq. (30) becomes

$$\dot{\mathbf{a}}_2(t) = -\Gamma(\mathbf{a}(t)) \cdot \gamma(\mathbf{a}(t)) \tag{32}$$

where

$$\Gamma(\mathbf{a}) = \int_0^\infty ds \lim_{\xi \rightarrow 0} \{ \xi^{-2} \langle [e^{Ls} \dot{\mathbf{A}}(\mathbf{X})] \Delta \dot{\mathbf{A}}(\mathbf{X}) \rangle \} \tag{33}$$

and  $\mathbf{v}_2$  has been neglected as promised. Alternatively,  $\Gamma$  may be written in the more compact notation as

$$\Gamma(\mathbf{a}(t)) = \int_0^\infty ds \lim_{\xi \rightarrow 0} [ \xi^{-2} \langle \dot{\mathbf{A}}(t+s) \Delta \dot{\mathbf{A}}(t) \rangle_t ] \tag{34}$$

The time-local nonlinear transport equations to order  $\xi^2$  are obtained by combining Eqs. (28), (29), and (32). The result is

$$\dot{\mathbf{a}}(t) = \xi \lim_{\xi \rightarrow 0} [ \xi^{-1} \mathbf{v}(\mathbf{a}(t)) ] - \xi^2 \Gamma(\mathbf{a}(t)) \cdot \gamma(\mathbf{a}(t)) \tag{35}$$

where the nonlinear transport coefficient matrix is given by Eq. (33). Equations (33) and (35) constitute our main result. They represent an important generalization of the more familiar linear transport theories, and yet are no more difficult to derive. These equations are valid far from equilibrium, and the transport coefficients accordingly manifest a nonlinear dependence upon the macroscopic state of the system.

We note that Eq. (33) has the same formal structure as the familiar time-correlation-function expressions for linear transport coefficients near equilibrium.<sup>(12)</sup> Here, however, the statistical average is weighted by  $\rho_0(\mathbf{X}; \mathbf{a})$  instead of the equilibrium distribution, and the quantity  $\Delta\dot{\mathbf{A}}$  is defined with reference to state  $\mathbf{a}$  instead of the equilibrium state. It is just these differences in weighting and  $\Delta\dot{\mathbf{A}}$  that make the transport coefficients state-dependent. This is a pleasingly intuitive result, which one might have expected to be a useful approximation, but which is actually exact in the time-local description to order  $\xi^2$ .

The time correlation function in Eq. (33) contains a "subtracted flux"<sup>(8)</sup> in the second factor only. If desired, however, a similar subtraction can be introduced into the first factor as well. One readily verifies that this correlation function is unchanged if  $e^{Ls}\dot{\mathbf{A}}(\mathbf{X})$  is replaced by  $e^{Ls}\dot{\mathbf{A}}(\mathbf{X}) - \delta\mathbf{A}(\mathbf{X}) \cdot \mathbf{M}(\mathbf{a}) \cdot \langle \delta\mathbf{A}\mathbf{G} \rangle$ , where  $\mathbf{G}(\mathbf{X})$  is arbitrary. The natural choice for  $\mathbf{G}$  is  $e^{Ls}\dot{\mathbf{A}}$ .

The transport coefficient matrix  $\Gamma(\mathbf{a})$  is not in general symmetric. However, it may be resolved into symmetric and antisymmetric parts in the usual way. All irreversible effects reside in the symmetric part, because the antisymmetric part does not contribute to the production of entropy. This is easily seen by combining Eqs. (14) and (35) to obtain

$$\dot{S}(\mathbf{a}) = k_B \xi^2 \gamma(\mathbf{a}) \cdot \Gamma(\mathbf{a}) \cdot \gamma(\mathbf{a}) \quad (36)$$

where use has been made of Eq. (10). The antisymmetric part of  $\Gamma(\mathbf{a})$  clearly drops out of  $\dot{S}(\mathbf{a})$ , and it therefore represents a reversible contribution of order  $\xi^2$  to  $\dot{\mathbf{a}}(t)$ . But this contribution, like that of  $\mathbf{v}_2$ , should ordinarily be negligible compared to the reversible terms of order  $\xi$ . To neglect it, one simply replaces  $\Gamma(\mathbf{a})$  in Eq. (35) by  $\frac{1}{2}[\Gamma(\mathbf{a}) + \Gamma^T(\mathbf{a})]$ , where superscript  $T$  denotes the transpose. In this approximation the matrix of transport coefficients becomes symmetric; one then has Onsager reciprocal relations far from equilibrium (cf. Hurley and Garrod,<sup>(16)</sup> García-Colín and del Río-Correa,<sup>(17)</sup> and Nettleton<sup>(18)</sup>).

The time-local transport equations contain quantities that involve limits as  $\xi \rightarrow 0$ . It should be noted that in some problems, the physical significance of  $\xi$  may require that the volume of the system be sent to infinity before  $\xi$  is sent to zero.

No attempt will be made here to carry the expansion beyond order  $\xi^2$ , although this can in principle be done. However, higher order terms in  $\xi$  are less likely to exhibit the behavior necessary for the transport equations to reduce to a time-local form. The physical reason for this is that the separation of time scales becomes less and less complete as  $\xi$  is increased, until eventually a point is reached at which the true nonlocal time behavior can no longer be consistently approximated by time-local equations. In hydrodynamics, for example, the divergence of the Burnett coefficients<sup>(4,9)</sup> indicates that a time-local description in terms of the conventional hydrodynamic variables breaks down at order  $\xi^3$ , and as already noted, it breaks down at order  $\xi^2$  in two dimensions. Attempts to push the expansion beyond second order are therefore hazardous and should be pursued with extreme caution. When time-local behavior does break down, it may be possible to restore it by including additional variables in the set  $\mathbf{a}$ . Unfortunately, it is rarely obvious what the additional variables should be.

### 5. LINEARIZATION ABOUT EQUILIBRIUM

Here we wish to derive the linearized transport equations to which the nonlinear equations of the previous section reduce when the system is sufficiently close to equilibrium.

In order for the general formalism to include equilibrium as a special case, the Hamiltonian  $H(\mathbf{X})$  must be a linear combination of the variables  $\mathbf{A}(\mathbf{X})$ .<sup>(8)</sup> That is,

$$H(\mathbf{X}) = \mathbf{h} \cdot \mathbf{A}(\mathbf{X}) \tag{37}$$

where  $\mathbf{h}$  is the constant vector whose components are the coefficients in the linear combination.

The equilibrium values of the variables  $\mathbf{a}$ , denoted by  $\mathbf{a}_{\text{eq}}$ , are determined by the condition

$$\gamma(\mathbf{a}_{\text{eq}}) = -\beta \mathbf{h} \tag{38}$$

where  $\beta = 1/k_B T$  and  $T$  is the equilibrium temperature. It then follows from Eqs. (4) and (5) that  $\rho_0(\mathbf{X}; \mathbf{a}_{\text{eq}}) = \rho_{\text{eq}}(\mathbf{X})$ , where

$$\rho_{\text{eq}}(\mathbf{X}) = \frac{\exp[-\beta H(\mathbf{X})]}{\int d\mathbf{X} \exp[-\beta H(\mathbf{X})]} \tag{39}$$

is the usual canonical equilibrium distribution. The equivalence of Eq. (38) to the explicit formula  $\mathbf{a}_{\text{eq}} = \int d\mathbf{X} \rho_{\text{eq}}(\mathbf{X}) \mathbf{A}(\mathbf{X})$  follows from Eq. (6).

Since  $H(\mathbf{X})$  is a constant of the motion,  $\dot{H} = \mathbf{h} \cdot \dot{\mathbf{A}} = 0$ . Equation (38) then implies the useful relation

$$\dot{\mathbf{A}}(\mathbf{X}) \cdot \boldsymbol{\gamma}(\mathbf{a}_{\text{eq}}) = 0 \tag{40}$$

We are now in a position to linearize Eq. (35). This is done by letting  $\mathbf{a}(t) = \mathbf{a}_{\text{eq}} + \delta\mathbf{a}(t)$ , expanding in powers of  $\delta\mathbf{a}$ , and discarding terms of order  $(\delta\mathbf{a})^2$  and higher. When this is done, all terms involving  $\boldsymbol{\gamma}(\mathbf{a}_{\text{eq}})$  drop out because of Eq. (40), and we are left with

$$\delta\dot{\mathbf{a}}(t) = -[\xi \lim_{\xi \rightarrow 0} (\xi^{-1} \langle \delta\mathbf{A}\dot{\mathbf{A}} \rangle_{\text{eq}}) + \xi^2 \boldsymbol{\Gamma}(\mathbf{a}_{\text{eq}})] \cdot \mathbf{M}(\mathbf{a}_{\text{eq}}) \cdot \delta\mathbf{a}(t) \tag{41}$$

where use has been made of Eqs. (11) and (21), and the subscript “eq” implies evaluation at  $\mathbf{a} = \mathbf{a}_{\text{eq}}$ . These are indeed just the usual equations of linear transport theory<sup>(12)</sup> (appropriately generalized to the case of nonzero  $\langle \delta\mathbf{A}\dot{\mathbf{A}} \rangle_{\text{eq}}$ ). The transport coefficients therein are now evaluated at  $\mathbf{a} = \mathbf{a}_{\text{eq}}$ . The average in Eq. (33) thus becomes an equilibrium average, whereupon Eq. (33) reduces to the usual expression for the linear transport coefficients in terms of equilibrium time correlation functions<sup>(12)</sup> (again for nonzero  $\langle \delta\mathbf{A}\dot{\mathbf{A}} \rangle_{\text{eq}}$ ).

## 6. NONLINEAR FRICTION COEFFICIENT OF A HEAVY PARTICLE

As an illustrative application of the preceding general formalism, we analyze the nonlinear momentum relaxation of a heavy particle of mass  $M$  immersed in a bath of  $N$  light particles of mass  $m$ . The microscopic state  $\mathbf{X}$  consists of the positions  $\mathbf{r}^N$  and momenta  $\mathbf{p}^N$  of the light particles, together with the position  $\mathbf{R}$  and momentum  $\mathbf{P}$  of the heavy particle. The set of macroscopic variables is taken to be  $\mathbf{A} = (\mathbf{U}, H)$ , where  $\mathbf{U} = \mathbf{P}/M$  is the velocity of the heavy particle and

$$H = |\mathbf{P}|^2/2M + \sum_i |\mathbf{p}_i|^2/2m + \Phi(\mathbf{r}^N, \mathbf{R}) \tag{42}$$

is the Hamiltonian of the entire system. Here  $\Phi(\mathbf{r}^N, \mathbf{R})$  is the total potential energy, including interactions between the heavy and light particles. The set of time derivative variables is simply  $\dot{\mathbf{A}} = M^{-1}(\mathbf{F}, 0)$ , where  $\mathbf{F} = \mathbf{F}(\mathbf{r}^N, \mathbf{R}) = -(\partial/\partial\mathbf{R})\Phi(\mathbf{r}^N, \mathbf{R})$  is the total force exerted by the light particles on the heavy particle. The smallness parameter is taken to be  $\xi = M^{-1}$ . The constrained equilibrium distribution is given by

$$\rho_0 = Q^{-1} \exp(-\beta H + \boldsymbol{\alpha} \cdot \mathbf{U}) \tag{43}$$

where

$$Q = \int d\mathbf{r}^N d\mathbf{R} d\mathbf{p}^N d\mathbf{P} \exp(-\beta H + \boldsymbol{\alpha} \cdot \mathbf{U}) \tag{44}$$

and the Lagrange multipliers  $\boldsymbol{\alpha}$  and  $\beta$  are functions of the mean heavy-particle velocity  $\mathbf{u} = \langle \mathbf{U} \rangle = \int d\mathbf{X} \rho \mathbf{U}$  and the mean total energy  $E = \langle H \rangle = \int d\mathbf{X} \rho H$ . The set of mean macroscopic variables is just  $\mathbf{a} = (\mathbf{u}, E)$ , and their thermodynamic conjugates are  $\boldsymbol{\gamma} = (\boldsymbol{\alpha}, -\beta)$ .

The momentum integrations in  $Q$  can be carried out explicitly, with the result

$$Q = (2\pi m/\beta)^{3N/2} (2\pi M/\beta)^{3/2} \exp(|\boldsymbol{\alpha}|^2/2\beta M) Z(\beta) \tag{45}$$

where

$$Z(\beta) = \int d\mathbf{r}^N d\mathbf{R} \exp[-\beta\Phi(\mathbf{r}^N, \mathbf{R})] \tag{46}$$

is the configurational partition function for the system. The functional dependence of  $\boldsymbol{\alpha}$  and  $\beta$  on  $\mathbf{u}$  and  $E$  is determined by Eq. (7), from which we obtain the relations

$$\boldsymbol{\alpha} = \beta M \mathbf{u} \tag{47}$$

$$\beta = f^{-1}(E - \frac{1}{2}M |\mathbf{u}|^2) \tag{48}$$

where the function  $f$  is defined by  $f(\beta) = 3(N+1)/2\beta - \partial[\ln Z(\beta)]/\partial\beta$ . Notice that  $f(\beta)$  is just the *equilibrium* energy of the system at a temperature of  $1/k_B\beta$ . It follows that  $1/k_B\beta$  is the temperature that the system would have at equilibrium if its energy were  $E - \frac{1}{2}M |\mathbf{u}|^2$ .

The streaming velocity is  $\mathbf{v} = M^{-1}(\langle \mathbf{F} \rangle, 0)$ , and clearly

$$\langle \mathbf{F} \rangle = Z^{-1} \int d\mathbf{r}^N d\mathbf{R} \exp[-\beta\Phi(\mathbf{r}^N, \mathbf{R})] \mathbf{F}(\mathbf{r}^N, \mathbf{R}) = 0 \tag{49}$$

The streaming velocity therefore vanishes in this problem. One readily verifies that the matrix  $\langle \delta \mathbf{A} \dot{\mathbf{A}} \rangle$  also vanishes in a similar manner, so that  $\Delta \dot{\mathbf{A}} = \dot{\mathbf{A}}$ . The matrix of transport coefficients therefore becomes

$$\boldsymbol{\Gamma} = \begin{pmatrix} \mathbf{g} & | & \mathbf{0} \\ \hline \mathbf{0} & | & 0 \end{pmatrix} \tag{50}$$

where  $\mathbf{0}$  is the zero vector and  $\mathbf{g}$  is a dyadic given by

$$\mathbf{g} = \int_0^\infty ds \lim_{M \rightarrow \infty} \langle (e^{Ls} \mathbf{F}) \mathbf{F} \rangle \tag{51}$$

The transport equations (35) now become

$$\dot{\mathbf{u}} = -(\beta/M)\mathbf{g} \cdot \mathbf{u} \quad (52)$$

$$\dot{E} = 0 \quad (53)$$

the latter of which simply expresses the conservation of total energy ( $\dot{H} = 0$ ). The velocity-dependent friction coefficient is the tensor  $\beta\mathbf{g}$ , which depends on  $\mathbf{u}$  through both  $\beta$  and  $\mathbf{g}$ . The dependence of  $\mathbf{g}$  on  $\mathbf{u}$  is due to the fact that the correlation function in Eq. (51) is to be evaluated by averaging over the constrained equilibrium distribution of Eq. (43). The combination of Eqs. (48), (51), and (52) generalizes the familiar expression for the linear friction coefficient<sup>(12)</sup> to the nonlinear regime.

Although the nonlinear friction coefficient is a tensor, it may be replaced by an equivalent scalar coefficient that depends on  $\mathbf{u}$  only through  $|\mathbf{u}|$ . The symmetry of the problem implies that  $\dot{\mathbf{u}}$  must be collinear with  $\mathbf{u}$ , so that  $\dot{\mathbf{u}} = (\hat{\mathbf{u}} \cdot \dot{\mathbf{u}})\hat{\mathbf{u}}$ , where  $\hat{\mathbf{u}} = \mathbf{u}/|\mathbf{u}|$ . Equation (52) may therefore be rewritten as

$$\dot{\mathbf{u}} = -(\beta/M)(\hat{\mathbf{u}} \cdot \mathbf{g} \cdot \hat{\mathbf{u}})\mathbf{u} \quad (54)$$

which shows that the effective scalar friction coefficient is simply

$$\beta(\hat{\mathbf{u}} \cdot \mathbf{g} \cdot \hat{\mathbf{u}}) = \beta \int_0^\infty ds \lim_{M \rightarrow \infty} \langle (e^{Ls}\mathbf{F} \cdot \hat{\mathbf{u}})\mathbf{F} \cdot \hat{\mathbf{u}} \rangle \quad (55)$$

## ACKNOWLEDGMENT

This work was performed under the auspices of the U.S. Department of Energy.

## REFERENCES

1. R. Zwanzig, *Ann. Rev. Phys. Chem.* **16**:67 (1965).
2. H. Mori *et al.*, eds., *Nonlinear Nonequilibrium Statistical Mechanics*, *Prog. Theo. Phys. Suppl.* **64** (1978).
3. L. Garrido, ed., *Systems Far from Equilibrium* (Lecture Notes in Physics, Vol. 132, Springer-Verlag, Berlin, 1980).
4. H. J. M. Hanley, ed., *Nonlinear Fluid Behavior*, *Physica* **118A** (1981).
5. P. Resibois and M. De Leener, *Classical Kinetic Theory of Fluids* (Wiley-Interscience, New York, 1977).
6. S. A. Rice and P. Gray, *The Statistical Mechanics of Simple Liquids* (Wiley-Interscience, New York, 1965), Sec. 7.3.D, and references cited therein.
7. J. W. Dufty, *Phys. Rev.* **176**:398 (1968).
8. B. Robertson, in *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus, eds. (MIT Press, Cambridge, Massachusetts, 1979), p. 289.



9. J. J. Brey, R. Zwanzig, and J. R. Dorfman, *Physica* **109A**:425 (1981), and references cited therein.
10. H. Grabert, *Projection Operator Techniques in Nonequilibrium Statistical Mechanics* (Springer-Verlag, Berlin, 1982).
11. D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin, Reading, Massachusetts, 1975).
12. R. Zwanzig, *J. Chem. Phys.* **40**:2527 (1964).
13. E. T. Jaynes, in *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus, eds. (MIT Press, Cambridge, Massachusetts, 1979), p. 15.
14. C. K. Wong, J. A. McLennan, M. Lindenfeld, and J. W. Dufty, *J. Chem. Phys.* **68**:1563 (1978).
15. J. W. Dufty and M. J. Lindenfeld, *J. Stat. Phys.* **20**:259 (1979).
16. J. Hurley and C. Garrod, *Phys. Rev. Lett.* **48**:1575 (1982).
17. L. S. García-Colín and J. L. del Río-Correa, *Phys. Rev. A* **30**:3314 (1984).
18. R. E. Nettleton, *Physica* **132A**:143 (1985).