

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION
SPONSORED PROJECT INITIATION

Date: February 29, 1980

Project Title: Fluctuation Phenomena In Thermal Physics

Project No: G-41-602

Project Director: Dr. R. F. Fox

Sponsor: National Science Foundation; Washington, D. C. 20550

Agreement Period: From January 15, 1980 Until June 30, 1982 (Grant Period)
Includes the usual 6 month unfunded flexibility period.

Type Agreement: Grant No. PHY-7921541

Amount: \$26,754 NSF
7,026 GIT (G-41-334)
\$33,780

Reports Required: Annual Technical Report; Final Project Report

Sponsor Contact Person (s):

Technical Matters

Dr. Richard A. Isaacson
Program Director for Theoretical Physics
Division of Physics
Directorate for Mathematical and Physical
Sciences and Engineering
National Science Foundation
Washington, D. C. 20550
(202) 632-4302

Contractual Matters

(thru OCA)

Mr. Philip M. King
Grants Official
National Science Foundation
Washington, D. C. 20550
(202) 632-5964

Defense Priority Rating: None

Assigned to: Physics (School/Laboratory)

COPIES TO:

Project Director
Division Chief (EES)
School/Laboratory Director
Dean/Director-EES
Accounting Office
Procurement Office
Security Coordinator (OCA)
Reports Coordinator (OCA)

Library, Technical Reports Section
EES Information Office
EES Reports & Procedures
Project File (OCA)
Project Code (GTRI)
Other OCA Research Property Coordinator

SPONSORED PROJECT TERMINATION/CLOSEOUT SHEET

Date 10/25/83

Project No. G-41-602

School ~~XXX~~ Physics

Includes Subproject No.(s) N/A

Project Director(s) R.F. Fox

~~GTRI~~

Sponsor NSF, Washington D.C.

Title: Fluctuation Phenomena In Thermal Physics

Effective Completion Date: 6/30/83 (Performance) 6/30/83 (Reports)

Grant/Contract Closeout Actions Remaining:

- None
- Final Invoice or Final Fiscal Report
- Closing Documents
- Final Report of Inventions
- Govt. Property Inventory & Related Certificate
- Classified Material Certificate
- Other _____

Continues Project No. _____

Continued by Project No. G-41-633

COPIES TO:

- Project Director
- Research Administrative Network
- Research Property Management
- Accounting
- Procurement/EES Supply Services
- Research Security Services
- ~~Reports Coordinator (OCA)~~
- Legal Services

- Library
- GTRI
- Research Communications (2)
- Project File
- Other _____

Progress in the broad area of fluctuation phenomena in thermal physics has been steady on several fronts during the past year. Three, more specific areas of effort are recognizable in this program, and are: 1) mathematical foundations of stochastic descriptions in physics; 2) applications of stochastic physics to problems in fluids and condensed matter; and 3) quantum mechanical fluctuation phenomena.

Last June, Ulrich Steiger earned his Ph.D. in physics under my supervision. We studied the Brownian motion of molecules in liquids with emphasis on coupled translational and rotational motion. While chemists have shown a rapidly increasing interest in such phenomena over the past two years, our interest was primarily formal. In "Coupled Translational and Rotational Diffusion in Liquids" (see attached publication list), we cleared up the earlier literature a bit, and developed the cumulant methods for such problems. In our more recent "Boson Operator Representation of Brownian Motion" we present a remarkable reformulation of classical Brownian motion, for both translation and rotation, in terms of boson creation and annihilation operators. This work provides still another example of the strong parallel between field theory methods and statistical mechanics methods. We expect to be able to clearly explain orientational effects in diffusion controlled reactions using these methods. Some progress in collaboration with Dr. J. Keizer at the University of California at Davis, with whom Dr. Steiger is currently doing post-doctoral research, has already been made in this direction.

The study of hydrodynamic fluctuations continues earlier work in this area. With Ph.D. candidate Byron Burel, light scattering from fluids out of equilibrium is being studied. Progress has been made on the effects of a thermal gradient on the diffusion of solute species in a fluid. This problem

is a variation on the more widely studied problem of light scattering from the fluid itself. The diffusion problem provides a somewhat more tractable example and we have been able to include vessel boundary condition effects as well as find the eigenfunction solutions analytically. It remains to definitively verify the theoretical approach for non-equilibrium processes. A paper reviewing the situation is submitted for publication by request of the editor of the Journal of Physical Chemistry. Another aspect of our work in hydrodynamical fluctuations deals with the non-linear terms in the theory. There is some controversy regarding the stress-strain non-linearity which was previously discussed in our earlier work. Recently, this work was criticized by Saarloos, Bedeaux, and Mazur. However, we found criticisms of their criticisms and have presented a paper elucidating the situation. This paper will soon appear in Physica A, where the issue was most recently raised.

During the last half year or so, our primary effort has been in the area of long time tails in autocorrelation functions. While developing an alternative to the Mori-Zwanzig theory for transport processes, the long time tail issue was directly confronted. Our approach, based on cumulant methods provided a time-convolutionless theory, which is equivalent to the Mori-Zwanzig theory, but which is formally more attractive and tractable. In particular, the 2nd cumulant directly provides the Green-Kubo formula for the transport coefficients, provided that corresponding autocorrelation functions are sufficiently well-behaved. The issue is whether or not long time tails in these correlations are "well"-behaved. Our study of such tails for diffusion has involved the mode coupling method, the Stokes-Boussinesq method, and the rigorous method of Sinai. We are finding that: 1) the theory is not as sound as has been claimed; and 2) if the long time tails are indeed really present, their effect is very small (in 3 dimensions). Although we can not yet vindicate the cumulant approach

definitively, we hope we will soon be preparing a manuscript on this complex and important question.

Quantum mechanical contexts for the study of stochastic phenomena have been a central part of this research program for a decade. In recent years, our efforts here had diminished, largely because experiments of relevance to our efforts were scarce. Now, however, with the development of "noisy" lasers, this work has revived. We are presently putting more effort here again in order to study the behavior of molecules in "noisy" laser beams. Such studies relate both to the fundamental theory of stochastic quantum mechanics and to applications to energy (light) driven molecular processes.

Attachment: Publication List

- 1) The Origin of Irreversibility in Quantum Statistical Mechanics, to appear in Advances in Mathematics (by invitation).
- 2) Coupled Translational and Rotational Diffusion in Liquids, J. Math. Phys. 23, 296 (1982).
- 3) Boson Operator Representation of Brownian Motion, accepted to appear in J. Math. Phys. (1982)
- 4) The Ideal Gas and the Second Law of Thermodynamics, accepted to appear in Amer. J. of Phys. (1982).
- 5) Stress-Strain Fluctuations in Non-Linear Hydrodynamics, accepted to appear in Physica A (1982).
- 6) Testing Theories of Nonequilibrium Processes with Light Scattering Techniques, invited for Journal of Phys. Chem.

Annual Report by Ronald F. Fox for NSF Grant PHY-7921541:

"Fluctuation Phenomena in Thermal Physics"

This report covers the second year of activity under this grant, and serves to also justify funding for a third year (1982) in accordance with the "Continuing Research Grant" status of the original award.

In June, 1981, my graduate student, Ulrich Steiger, received his Ph.D. degree. His thesis: "Analysis of Coupled Translational and Rotational Diffusion Using Operator Calculus" culminated a program of research started seven years ago by me. Mr. Steiger was able to bring this program to fruition in just two years. Two major results were achieved. 1) The contraction of the Kramers-Liouville equation into the Smoluchowski equation was obtained in a very general setting, and was applied to the complicated phenomenon of coupled translational and rotational diffusion. 2) A boson operator representation for the Kramers-Liouville operator was discovered which greatly facilitates the analysis, and illustrates deeper connections with field theory. The first of several papers based on the thesis has been accepted for publication in the Journal of Mathematical Physics.

Another graduate student, Byron Burel, has made much progress on the problem of light scattering from fluids in nonequilibrium states. Several research groups have published work on this problem for light scattering from H_2O subject to a thermal gradient. Modifications of the Brillouin doublet are predicted. The problem has generated several papers which were later corrected, and concurrence on the "correct" result was slow to come. We analyzed this problem thoroughly and verified the current correct result. Moreover, we discovered the reason behind a perplexing cancellation of first order perturbation results in all of the earlier work. By looking at a somewhat more

simple problem, diffusion in a fluid in a thermal gradient, we achieved an exactly solvable case, for which perturbation techniques are unnecessary. A simple symmetry argument proves to be the basis for the cancellation in the perturbation problem. Our review of the earlier work will appear in Accounts of Chemical Research, by invitation from the Editor, J. F. Burnett. An account of our exactly solvable case is in preparation for Physical Review.

During the Spring quarter, 1981, I was Visiting Professor of Chemistry at the University of California at Davis. There, I enjoyed a collaboration with Drs. Joel Keizer in Chemistry and Marc Mangel in Mathematics, both of whom share my research interests in stochastic processes and nonequilibrium statistic mechanics. On May 7, 1981, Dr. I. Prigogine delivered the Hitchcock lecture, as a guest of the Department of Chemistry at Davis. He and I met for the first time and spent 90 minutes together with a small group put together by Dr. D. McQuarrie. Prigogine told me personally that he was mistaken in his debate with me over the Glandsdorff-Prigogine stability criterion. Even though he will not publically declare his error, he has admitted it to me privately! After seven years of debate, this was a most welcome conclusion.

In March, 1981, a paper appeared in Physical Review D in which two Georgia Tech colleagues and I put straight a mistaken claim by T. H. Boyer that the Boltzmann distribution and the Planck spectrum were inconsistent with each other for relativistic particles. What is remarkable here is that we submitted our paper in September, 1979. The 18 month delay until the paper appeared was a result of repeated rejections and resubmittals which finally terminated when Boyer finally saw the light. The entire issue caught someone's attention at the Scientific American where the debate was featured under "Science and the Citizen" in the June, 1981 issue. Currently, another graduate student, A. Teate, is working on related problems for multiphoton and multiphonon processes.

The Advances in Mathematics editor, G-C. Rota, invited a paper to celebrate Mark Kac's 65th birthday. I submitted a paper entitled "The Origin of Irreversibility in Quantum Statistical Mechanics" which reviews my work with reduced density matrices and cumulants for operators. A publication date is still not set.

ANALYSIS OF COUPLED TRANSLATIONAL AND ROTATIONAL
DIFFUSION USING OPERATOR CALCULUS

A THESIS

Presented to
The Faculty of the Division
of Graduate Studies

By
Ulrich Robert Steiger

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the School of Physics

Georgia Institute of Technology

June 1981

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Ronald F. Fox, for proposing this thesis topic and teaching me this area of physics, and for all the valuable suggestions and clarifying discussions which made the work a pleasure. I would like to thank Dr. Joseph Ford and Dr. Theodore P. Hill for many stimulating discussions and for reading the text, and Ms. Gail Tucker for the excellent typing of this manuscript. However, most of all, I am grateful to my parents for their understanding and support.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	ii
SUMMARY	v
Chapter	
I. OPERATOR CALCULUS AND TRANSLATIONAL DIFFUSION	1
II. COUPLED TRANSLATIONAL AND ROTATIONAL DIFFUSION	15
Kramers-Liouville Equation	
Scalar Products and Partial Differential Operators	
The Operators A , $A^\#$, B and $B^\#$	
Remarks on Contraction Semi-Groups Generated by the Kramers-Liouville Operator	
The Liouville Operator in the Interaction Picture	
The Moments $\langle L(t_1) \dots L(t_n) \rangle$	
The Diffusion Equation (Second Cumulant)	
N-Particle Diffusion	
III. CORRECTIONS OF THE TRANSLATIONAL SMOLUCHOWSKI EQUATION	66
Gaussian Property of the Liouville Operator in the Interaction Picture (Translational Motion)	
Generalized U-W Relations	
The Fourth Cumulant: Position Dependent Diffusion Coefficients	
The Sixth Cumulant: A Fourth Order Partial Differential Operator	
Remarks on the Structure of the Complete Cumulant Expansion	
IV. EXAMPLES AND APPLICATIONS	103
Axially Symmetric Particles	
Diffusion in Two Dimensions	
The Harmonic Oscillator	
Position Dependent Diffusion Coefficients	
V. CONCLUSIONS AND RECOMMENDATIONS	127
APPENDIX I	135

	Page
APPENDIX II	144
BIBLIOGRAPHY	148
VITA	150

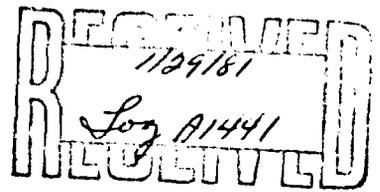
SUMMARY

The equations for coupled translational and rotational diffusion of asymmetric molecules immersed in a fluid are obtained. The method used begins with the Kramers-Liouville equation and leads to the generalized Smoluchowski equation for diffusion in the presence of potentials. Both external potentials and intermolecular potentials are considered. The contraction of the description from the Kramers-Liouville equation to the Smoluchowski equation is achieved by using a combination of operator calculus and cumulants. Explicit solutions of these equations are obtained in the two-dimensional case. The formalism also allows the calculation of corrections to the generalized Smoluchowski equation. Smoluchowski's result is precisely the second cumulant, in the cumulant expansion.

The next non-vanishing term, the fourth cumulant, leads to diffusion equations with position dependent diffusion coefficients. The higher order cumulants lead to evolution equations for the reduced probability density which contain partial derivatives of order m with $m \geq 3$. Explicit expressions are given up to the sixth order in the cumulant expansion for translation diffusion. From a practical point of view, this formalism is very useful because partial differential equations can be solved numerically by using a finite element calculation.

The contraction of the Liouville-Kramers description into the Smoluchowski description is achieved by using a creation-destruction operator representation. In an appropriately defined inner product

space, these operators possess the operator algebra of boson operators in quantum field theory. The discovery of this representation for the Liouville-Kramers description including rotational effects greatly facilitated obtaining the Smoluchowski contraction.



Coupled Translational and Rotational

Diffusion in Liquids

Ulrich Steiger and Ronald F. Fox

School of Physics

Georgia Institute of Technology

Atlanta, GA 30332

December 1980

To appear in Journal of Mathematical Physics

ABSTRACT

The equations for coupled translational and rotational diffusion of asymmetric molecules immersed in a fluid are obtained. The method used begins with the Kramers-Liouville equation and leads to the generalized Smoluchowski equation for diffusion in the presence of potentials. Both external potentials and intermolecular potentials are considered. The contraction of the description from the Kramers-Liouville equation to the Smoluchowski equation is achieved by using a combination of operator calculus and cumulants. Explicit solutions to these equations are obtained for the two dimensional case. Comparison of our results with earlier literature is also presented.

I. Introduction

In this paper we study the translational and rotational motion of molecules immersed in a fluid. The molecules experience translational and rotational Brownian motion as a result of the bombardment by fluid molecules. The description of this essentially stochastic process in terms of the probability-distribution function $P(t,x)$ leads to a diffusion equation

$$\begin{aligned} \frac{\partial}{\partial t} P(t,x) &= \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} P(t,x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} P(t,x) \\ &\equiv A P(t,x) \end{aligned} \tag{1}$$

for all times $t \geq 0$ and all points x , $x = (q_1, q_2, q_3, \phi, \theta, \psi)$. $q = (q_1, q_2, q_3)$ describe the position and the Euler angles $\alpha = (\phi, \theta, \psi)$ fix the orientation. The differential operator A is a diffusion operator. All eigenvalues of the symmetric matrix $(a_{ij}(x))$ are nonnegative. For translational diffusion A_T is simply a diffusion constant multiplied by the Laplace operator. Favro [5] derived the diffusion equation for rotational Brownian motion and was able to solve it for axial symmetric molecules using the fact that the diffusion operator A_R has the same form as the quantum mechanical Hamilton operator for a rigid body [10], the properties of which are well known. In general the translational and rotational motions are coupled in a complicated way.

Already 50 years ago, Kolmogorov showed that under very general conditions a Markov process defined in terms of the transition probability $F(t,x,x')dx'$ of finding a particle initially at point x in the infinitesimal small set dx' after a lapse of time t , leads to a diffusion equation. The probability density

$$P(t,x) \equiv \int_{S_{x'}} F(t,x',x) P(0,x') dx' \tag{2}$$

satisfies equation (1). S_x is the space containing all points x .

$P(0,x)$ is the initial distribution at time $t=0$.

The concept of a Markov process is an idealization of the underlying physical reality. For a complete dynamical description, it is necessary to consider the distribution function $f_c(t, x_c, y_c)$ defined on the phase space $S_{x_c} \times S_{y_c}$ consisting of all points (x_c, y_c) with $x_c = (q_1, q_2, q_3, \phi, \theta, \psi)$ and the canonically conjugate momenta $y_c = (p_1, p_2, p_3, p_\phi, p_\theta, p_\psi)$. The distribution function $f_c(t, x_c, y_c)$ satisfies the Kramers-Liouville equation [14], [15]

$$\frac{\partial}{\partial t} f_c(t, x_c, y_c) = (L+K) f_c(t, x_c, y_c) . \quad (3)$$

L is Liouville's operator and K denotes Kramers operator, which describes the effect of all random forces acting on the Brownian particle. If equation (3) can be solved for some initial distribution $f_c(0, x_c, y_c)$ then it is possible to find an operator $G(t, x_c)$ such that the averaged distribution $p(t, x_c)$ defined by

$$p(t, x_c) \equiv \int_{S_{y_c}} dy_c f(t, x_c, y_c) \quad (4)$$

fulfills the first order differential equation in time:

$$\frac{\partial}{\partial t} P(t, x_c) = G(t, x_c) P(t, x_c) . \quad (5)$$

In general nothing is gained, since $G(t, x_c)$ might be a very complicated operator. We will use the cumulant expansion [1], [2] to approximate the operator $G(t, x_c)$.

$$G(t, x_c) = \sum_{n=1}^{\infty} G^{(n)}(t, x_c) \quad (6)$$

It turns out, that the diffusion operator A is the first nonvanishing term in the expansion (6). Equation (1), where A is now replaced by the second cumulant $G^{(2)}(t, x_c)$, ($G^{(1)}(t, x_c) = 0$), is a very good approximation of (5).

K describes the time evolution of the distribution of the momenta due to random forces. The momenta $y_c(t)$ can be considered as random variables, which very quickly become independent. $y_c(t)$ is independent of $y_c(t + \Delta t)$ if the lapse of time Δt is large compared with the correlation time τ_k . It can be shown [3], that the n^{th} cumulant is proportional to

$$G^{(n)} \sim \hat{\tau}^{* n-1} \tau \quad (7)$$

$\hat{\tau}$ is a dimensionless quantity. $\hat{\tau} \equiv \tau_k / \tau$. τ is some typical macroscopic time unit.

Intuitively, it is clear that we obtain a Markov process on S_x described by (1) if the correlation time τ_k of the momenta $y_c(t)$ becomes very small. It is the short correlation time which makes the higher order contributions small.

The idea of deriving the diffusion operator A as the lowest order of a cumulant expansion (6) is not new. The actual calculation of the operators $A, G^{(3)}, \dots$ is complicated by the nonlinearity of the equation of motion for a

rigid body. The time derivatives of the angular momentum L' and translational momentum p' expressed in an orthogonal coordinate frame attached to the moving particle are:

$$\begin{aligned}\dot{L}' &= L' \times I^{-1} L' + N' \\ \dot{p}' &= p' \times I^{-1} L' + F'\end{aligned}\tag{8}$$

N' and F' are the torques and the forces acting on the particle. The prime denotes vectors in the body fixed coordinate frame. I is the tensor of inertia. It is necessary, to choose body fixed coordinates for both L' and p' since otherwise the friction tensor C depends on the orientation, (see (70)).

The purpose of this work is to analyse the rotational and translational diffusion in the most general case using a mathematically transparent method. We will show that:

(i) The generalized Smoluchowski equation is the lowest order contribution of $G(t, x_c)$. Starting off with a Maxwell distribution at time $t=0$ the diffusion tensor is time dependent. For $t < \tau_k$ the diffusion tensor depends on the mass and the moments of inertia, and becomes stationary for $t \gg \tau_k$.

(ii) The diffusion equation couples the translational and rotational degrees of freedom even in the simplest case. As an illustration, the two dimensional diffusion equation is solved. The solutions are obtained in terms of exponential and Mathieu functions. (Section V).

(iii) A suspension of N interacting Brownian particles leads to a diffusion equation for the N particle density $P(t, x_c^{(1)}, x_c^{(2)}, \dots, x_c^{(N)})$. (Section IV).

In Section II the operator calculus used later is introduced and applied to the translational motion. Section III treats coupled translational and rotational diffusion.

II. Operator Calculus, Translational Diffusion

The starting point of the theory is the Kramers-Liouville equation [14], [15].

$$\frac{\partial}{\partial t} f(t, q, p) = B f(t, q, p) = (L+K) f(t, q, p). \quad (9)$$

q are the coordinates describing the position, $q = (q_1, q_2, q_3)$ and p are the conjugate momenta. Liouville's operator is

$$L f = - m^{-1} p \cdot \frac{\partial}{\partial q} f + \frac{\partial U}{\partial q} \cdot \frac{\partial}{\partial p} f. \quad (10)$$

U denotes the potential. Kramers operator is

$$K f = \alpha \frac{\partial}{\partial p} \cdot (m^{-1} p + kT \frac{\partial}{\partial p}) f. \quad (11)$$

It is convenient [14] to work in the "interaction picture"

$$f \equiv e^{tK} \tilde{f} \quad (12)$$

The exponential e^{tK} is defined by a formal power series in tK and acts on the new function \tilde{f} which is assumed to be smooth enough, such that the series $e^{tK} \tilde{f} \equiv \sum_{n=0}^{\infty} \left((tK)^n / n! \right) \tilde{f}$ converges. The smoother \tilde{f} the smaller the contribution of $(tK)^n$ which is a differential operator of order $2n$ in the variable p .

The time evolution for \tilde{f} is governed by the Kramers-Liouville equation in the "interaction picture".

$$\frac{\partial}{\partial t} \tilde{f} = e^{-tK} L e^{tK} \tilde{f} \equiv \tilde{L}(t) \tilde{f} \quad (13)$$

The operator $\tilde{L}(t)$ can be expressed in terms of the differential operators $\frac{\partial}{\partial q}$ and $\frac{\partial}{\partial p}$ using the identity

$$e^{-tK} L e^{tK} = e^{-[K, \cdot]t} L \quad (14)$$

The proof of this equation is found in [1]. The operator on the right hand side is by definition

$$e^{-[K, \cdot]t} L \equiv L + \sum_{n=1}^{\infty} [K, \cdot]^n L \left(\frac{(-t)^n}{n!} \right) \quad (15)$$

The commutators $[K, \cdot]^n L$ can be defined by recursion

$$\begin{aligned} [K, \cdot]^1 L &\equiv [K, L] \\ [K, \cdot]^2 L &\equiv [K, [K, L]] \\ [K, \cdot]^n L &\equiv [K, \cdot]([K, \cdot]^{n-1} L) \end{aligned} \quad (16)$$

We can calculate all terms in the infinite sum (15). Applying the commutator algebra discussed in [14] leads to

$$\begin{aligned} \tilde{L}(t) &= -e^{-(\alpha/m)t} \frac{\partial}{\partial q} \cdot \left(\frac{p}{m} + kT \frac{\partial}{\partial p} \right) \\ &+ e^{(\alpha/m)t} \frac{\partial}{\partial p} \cdot \left(\frac{\partial U}{\partial q} + kT \frac{\partial}{\partial q} \right) \end{aligned} \quad (17)$$

In Section III the corresponding expression for translational and rotational motion is derived in great detail.

Formally, the solution of (17) can be written

$$\tilde{f}(t) = E(t) \tilde{f}_0 \equiv \mathbb{T} \exp \int_0^t ds \tilde{L}(s) \tilde{f}_0 \quad (18)$$

In which T_{\pm} exp is the time ordered exponential [1]. f_0 is the initial distribution. The time ordered exponential must be used because $L(t_1)$ does not commute with $L(t_2)$ if $t_1 \neq t_2$. We would like to derive the time evolution for the averaged distribution $P(t, q)$

$$\begin{aligned}
 P(t, q) &\equiv \int d^3 p f(t, q, p) = \int d^3 p e^{tK} \tilde{f}(t, q, p) \\
 &= \int d^3 p \tilde{f}(t, q, p) \equiv \langle \tilde{f}(t, q) \rangle
 \end{aligned}
 \tag{19}$$

The third equality can be proved by expanding the exponential e^{tK} . After integrating by parts, all but the lowest order term, which is \tilde{f} , vanish. We can assume that $\tilde{f}(t, q, p)|_{p_i=\infty} = 0$.

We write the initial condition $\tilde{f}(0, q, p) \equiv \tilde{f}_0(q, p) = f_0(q, p)$ in the form

$$f_0(q, p) = g(q, p) P_0(q)$$

$$P_0(q) = \langle f_0(q) \rangle$$

With equation (18), (19) and (20) one obtains

$$\begin{aligned}
 P(t, q) &= \int d^3 p \tilde{f}(t, q, p) \\
 &= \int d^3 p E(t) g(q, p) P_0(q) \\
 &\equiv \langle E(t) \rangle_g P_0(q)
 \end{aligned}
 \tag{21}$$

The operator $\langle E(t) \rangle_g$ is obtained by multiplying $g(q, p)$ from the left with $E(t)$ and integrating over the momenta p . Differentiating equation (21) with

respect to t gives the time evolution equation

$$\frac{\partial}{\partial t} P(t, q) = \left[\frac{\partial}{\partial t} \langle E(t) \rangle_g \right] \langle E(t) \rangle_g^{-1} P(t, q). \quad (22)$$

We expect, that the inverse $\langle E(t) \rangle_g^{-1}$ exists at least for small times. It may be obtained by the Neumann series [4] $A^{-1} = \sum_{n=0}^{\infty} (1-A)^n$. The operator

$$G(t, q) \equiv \left[\frac{\partial}{\partial t} \langle E(t) \rangle_g \right] \langle E(t) \rangle_g^{-1} \quad (23)$$

$$\frac{\partial}{\partial t} P(t, q) = G(t, q) P(t, q)$$

depends on q since $g(q, p)$ is a function on q and p . But in most physical applications the initial distribution of the momenta does not depend on the position q . In this case the operator G depends only on t .

In order to calculate $G(t)$ we use the cumulant expansion, [1], [2], [3], which is obtained by reordering the expression

$$G(t) = \sum_{n=0}^{\infty} \langle \tilde{L}(t) \prod_{+} \exp \int_0^t \tilde{L}(s) ds \rangle_g \langle 1 - \prod_{+} \exp \int_0^t \tilde{L}(s) ds \rangle_g^{-n} \quad (24)$$

$$G(t) = \sum_{\ell=1}^{\infty} G^{(\ell)}$$

Compare (18), (22), (23). $G^{(\ell)}$ contains all terms of the sum in (24) which are of order ℓ in the operator $\tilde{L}(s)$. The two lowest order terms are:

$$\begin{aligned}
 G^{(1)}(t) &= \langle \tilde{L}(t) \rangle_g = \int d^3p \tilde{L}(t) g(p) \\
 G^{(2)}(t) &= \int_0^t ds \langle \tilde{L}(t) \tilde{L}(s) \rangle_g - \int_0^t ds \langle \tilde{L}(t) \rangle_g \langle \tilde{L}(s) \rangle_g \\
 &= \int_0^t ds \int d^3p \tilde{L}(t) \tilde{L}(s) g(p) \\
 &\quad - \int_0^t ds \int d^3p \tilde{L}(t) g(p) \int d^3p' \tilde{L}(s) g(p')
 \end{aligned}
 \tag{25}$$

The higher order terms are given in Section VI.

We assume, that the distribution in the momenta is initially a Maxwell distribution

$$g(p) = (2\pi mkT)^{-3/2} \exp\left(-p^2/2mkT\right).
 \tag{26}$$

In this case, it is easy to verify that the first cumulant $G^{(1)}(t)$ vanishes for all times $t \geq 0$. The second cumulant is

$$G^{(2)}(t) = \frac{kT}{\alpha} \frac{\partial}{\partial q} \cdot \left(\frac{1}{kT} \frac{\partial U}{\partial q} + \frac{\partial}{\partial q} \right) \left(1 - e^{-\frac{\alpha}{m} t} \right)
 \tag{27}$$

The time evolution equation (23) is, to second order in \tilde{L} , the Smoluchowski equation with time-dependent diffusion "constant"

$$\Lambda(t) = \frac{kT}{\alpha} \left(1 - e^{-\frac{\alpha}{m} t} \right)
 \tag{28}$$

$$\frac{\partial}{\partial t} P(t, q) = \frac{\partial}{\partial q} \cdot \Lambda(t) \left(\frac{1}{kT} \frac{\partial U}{\partial q} + \frac{\partial}{\partial q} \right) P(t, q) .$$

At $t=0$ the diffusion constant vanishes since by assumption the distribution in p was given by a symmetric function, the Maxwell distribution. After a short time of order m/α the particles start moving until finally the Boltzmann distribution is reached. In order to illustrate the meaning of the time-dependent diffusion constant $\Lambda(t)$ we calculate the first and second cumulant with the initial distribution $g(p) = \delta(p-p_0)$. All particles have the same momentum p_0 at $t=0$. In this case the first cumulant does not vanish:

$$G_{\delta}^{(1)} = -e^{-\frac{\alpha}{m}t} p_0 m^{-1} \cdot \frac{\partial}{\partial q} \quad (29)$$

$$G_{\delta}^{(2)}(t) = \frac{1}{\alpha} \left(e^{-\frac{\alpha}{m}t} - e^{-\frac{2\alpha}{m}t} \right) \times \left\{ \frac{1}{m} \left(\frac{\partial}{\partial q} \cdot p_0 \right)^2 - kT \left(\frac{\partial}{\partial q} \right)^2 \right\} + \frac{1}{\alpha} \left(1 - e^{-\frac{\alpha}{m}t} \right) \frac{\partial}{\partial q} \cdot \left(\frac{\partial U}{\partial q} + kT \frac{\partial}{\partial q} \right) \quad (30)$$

In the limit $t \rightarrow \infty$ both expressions (27) and (29)(30) agree, as they should. The operator $G(t)$ is independent of the initial condition for large times. The larger α/m , the faster $G(t)$ approaches the constant expression. For very large values of α/m the dynamics governed by (23) approaches a Markov process.

Formally the Markovian limit is obtained by first rescaling the time $\tau = \alpha^{-1}t$ and taking the limit $\alpha \rightarrow \infty$. In this limit all higher cumulants vanish since they are proportional to higher powers of $\frac{1}{\alpha}$.

III. Coupled Translational and Rotational Diffusion

We consider particles of arbitrary shape in a fluid. The friction forces depend on the orientation. We will describe a proper choice for the variables.

In [12] and [13] inconsistent definitions which lead to wrong results are used.

The position and orientation of each particle is determined by the six variables comprised in the six-tuple x .

$$x = (q_1, q_2, q_3, \phi, \theta, \psi) \quad (31)$$

O is an arbitrary origin and C the center of mass. q_1, q_2, q_3 are the coordinates of the vector OC in the laboratory frame where $\hat{e}_1, \hat{e}_2, \hat{e}_3$ are three arbitrary orthogonal vectors of length one such that $\hat{e}_1 \times \hat{e}_2 = \hat{e}_3$, et cyclic. It is convenient, to choose the Euler angles $\alpha = (\phi, \theta, \psi)$ to describe the orientation [16]. We will also use the body fixed coordinate frame $\hat{e}'_1, \hat{e}'_2, \hat{e}'_3$ such that the tensor of inertia I becomes diagonal. The components of the vector \hat{e}'_1 expressed in the laboratory fixed frame $\hat{e}_1, \hat{e}_2, \hat{e}_3$ are

$$\begin{aligned} (\hat{e}'_1)_l &\equiv R_{li}(\phi, \theta, \psi) \\ R(\phi, \theta, \psi) &\equiv \left(R_{li}(\phi, \theta, \psi) \right) \end{aligned} \quad (32)$$

The Euler angles are defined by

$$R(\phi, \theta, \psi) \equiv R_z(\phi) R_x(\theta) R_z(\psi). \quad (33)$$

$R_z(\phi)$ and $R_z(\psi)$ are counter-clockwise rotations of a vector about the \hat{e}_3 axis. $R_x(\theta)$ is a rotation about the \hat{e}_1 axis.

$$\begin{aligned}
 R_z(\phi) &= e^{\phi T_3} \\
 R_x(\theta) &= e^{\theta T_1} \\
 R_z(\psi) &= e^{\psi T_3}
 \end{aligned}
 \tag{34}$$

The 3 x 3 matrices T_1, T_2, T_3 are defined

$$(T_1)_{\ell m} = \epsilon_{1\ell m} \tag{35}$$

ϵ_{ilm} is the completely antisymmetric Levi-Civita tensor. Besides the position x (31) we need the momenta y

$$y = (p'_1, p'_2, p'_3, L'_1, L'_2, L'_3) \tag{36}$$

Both the translational momenta p' and the angular momenta L' are expressed in the body fixed coordinate frame. The tensor of inertia and the friction tensor depend only on the mass distribution and shape of the particle. They are independent of the orientation if body fixed coordinates are used. According to (32) the vector p' and $p \equiv m\dot{q}$, where m is the mass and the dot denotes the time derivative, are related in the following way:

$$\begin{aligned}
 p' &= R^\dagger(\phi, \theta, \psi)p \\
 &= R^{-1}(\phi, \theta, \psi)p
 \end{aligned}
 \tag{37}$$

The angular momentum L' is the product of the angular velocity ω' and the tensor of inertia I

$$L' = I \omega' \tag{38}$$

With equation (37) the skewsymmetric angular velocity matrix Ω [11] expressed in the body fixed frame is

$$\Omega = R^{-1} \dot{R} \quad (39)$$

The matrix Ω and the pseudovector ω' are related:

$$\Omega = \sum_{i=1}^3 \omega'_i T_i \quad (40)$$

In order to obtain Ω in terms of the Euler angles $\alpha = (\phi, \theta, \psi)$ and their time derivatives we substitute in (39) for the rotation R the expressions (33) and (34). Evaluating the time derivative in (39) and multiplying \dot{R} from the left with R^{-1} leads to

$$\begin{aligned} \Omega = & \dot{\phi} e^{-\psi T_3} e^{-\theta T_1} T_3 e^{\theta T_1} e^{\psi T_3} \\ & + \dot{\theta} e^{-\psi T_3} T_1 e^{\psi T_3} + \dot{\psi} T_3 \end{aligned} \quad (41)$$

We compare this expression with (40). Equation (41) can be simplified using the commutator algebra $[T_i, T_j] = \epsilon_{ijk} T_k$ [10], [16]. One obtains for the angular velocity ω'

$$\begin{aligned} \omega'_1 &= \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \\ \omega'_2 &= \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \\ \omega'_3 &= \dot{\psi} + \dot{\phi} \cos \theta \end{aligned} \quad (42)$$

Now we are able to describe the motion of the particle completely. The phase space $S_x \times S_y$ consists of all pairs $z = (x, y)$ defined by (31), (36), (37), (38) and (42).

Liouville's Equation

The motion of the rigid body is a solution of the canonical equations

$$\dot{x}_c = \frac{\partial H}{\partial y_c} \quad \dot{y}_c = - \frac{\partial H}{\partial x_c} \quad (43)$$

$$H(x_c, y_c) = \frac{1}{2m} ||p||^2 + \frac{1}{2} L' \cdot I^{-1} L' + U(x_c)$$

The canonical conjugate variables x_c and y_c are $x_c = x$ and $y_c = (p_1, p_2, p_3, p_\phi, p_\theta, p_\psi)$.

The canonical conjugate momenta for the angle variables $\alpha = (\phi, \theta, \psi)$ are given by

$$p_\alpha = \frac{\partial T}{\partial \dot{\alpha}} \quad \text{with } T \equiv \frac{1}{2} L' \cdot I^{-1} L'$$

$$p_\phi = L'_1 \sin \theta \sin \psi + L'_2 \sin \theta \cos \psi + L'_3 \cos \theta$$

$$p_\theta = L'_1 \cos \psi - L'_2 \sin \psi \quad (44)$$

$$p_\psi = L'_3$$

For every solution $z_c(t) \equiv (x_c(t), y_c(t))$ of equation (43) Liouville's theorem holds

$$\frac{\partial}{\partial t} f_c(t, z_c) + \dot{z}_c \cdot \frac{\partial}{\partial z_c} f_c(t, z_c) = 0 \quad (45)$$

It would be more convenient to express the particle density distribution f_c as a function of the variables $z = (x, y)$ defined earlier, instead of as a function of $z_c = (x_c, y_c)$. We define a new density

$$f(t, z) \equiv f_c(t, z_c(z)) \quad (46)$$

With the following identifies, one obtains the Liouville equation (48) for the new density $f(t,z)$.

$$\frac{\partial}{\partial z_c} = \frac{\partial z}{\partial z_c} \frac{\partial}{\partial z}$$

$$\dot{z} = \frac{d}{dt} z(t) \equiv \frac{d}{dt} z(z_c(t)) = \frac{\partial z}{\partial z_c} \dot{z}_c \quad (47)$$

$$\frac{\partial z}{\partial z_c} \frac{\partial z_c}{\partial z} = \mathbb{1}_{12}$$

$\mathbb{1}_{12}$ is the 12 dimensional identity matrix. We get:

$$\frac{\partial}{\partial t} f(t,z) + \dot{z} \frac{\partial}{\partial z} f(t,z) = 0 \quad (48)$$

The transformation $z_c = z_c(z)$ is given by equation (37) and (44). The Jacobian determinant is $-\sin \theta$. For any observable $O = O(z_c)$ the expectation value $EO \equiv \int dz_c O(z_c) f_c(t, z_c)$ can also be expressed in the new variables $z = (q, \alpha, p', L')$

$$\begin{aligned} EO &= \int dz \left| \text{Det} \left(\frac{\partial z_c}{\partial z} \right) \right| O(z_c(z)) f(t, z) \\ &= \int d^3 q d\phi d\theta \sin\theta d\psi d^3 p' d^3 L' \end{aligned} \quad (49)$$

$$\times O(q, \phi, \theta, \psi, p', L') f(t, q, \phi, \theta, \psi, p', L')$$

Equation (45) and (48) are formally the same but the meaning of the differential operators $\frac{\partial}{\partial z_c}$ and $\frac{\partial}{\partial z}$ are very different.

$$\frac{\partial}{\partial z_c} = \left(\frac{\partial}{\partial x_c}, \frac{\partial}{\partial y_c} \right), \quad \frac{\partial}{\partial z} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) \quad (50)$$

The gradient $\frac{\partial}{\partial x_c}$ is evaluated with the canonical conjugate momenta

$y_c = (p_1, p_2, p_3, p_\phi, p_\psi)$ fixed. When $\frac{\partial}{\partial x}$ operates, the momenta $y = (p_1', p_2', p_3', L_2', L_3')$ are fixed.

$$\frac{\partial}{\partial x_c} = \left(\left(\frac{\partial}{\partial q_1} \right)_{q_2, q_3, \phi, \theta, \psi}, \dots \right)_{p_1, p_2, p_3, p_\phi, p_\theta, p_\psi}, \quad \left(\frac{\partial}{\partial \psi} \right)_{q_1, q_2, q_3, \phi, \theta}, \dots \right)_{p_1, p_2, p_3, p_\phi, p_\theta, p_\psi}$$

$$\frac{\partial}{\partial x} = \left(\left(\frac{\partial}{\partial q_1} \right)_{q_2, q_3, \phi, \theta, \psi}, \dots \right)_{p_1', p_2', p_3', L_1', L_2', L_3'}, \quad \left(\frac{\partial}{\partial \psi} \right)_{q_1, q_2, q_3, \phi, \theta}, \dots \right)_{p_1', p_2', p_3', L_1', L_2', L_3'}$$

(51)

Rather than using (47) to calculate \dot{z} we go back to Euler's equation.

$$\frac{\partial}{\partial x_c} \mathcal{L} - \frac{d}{dt} \frac{\partial}{\partial \dot{x}_c} \mathcal{L} = 0 \quad (52)$$

The Lagrange function \mathcal{L} is $\mathcal{L} = \frac{1}{2} y^\dagger M^{-1} y - U(x_c)$. M is the generalized inertia matrix

$$M = \begin{pmatrix} m \mathbb{1}_3 & 0 \\ 0 & I \end{pmatrix} \quad (53)$$

M is a symmetric 6 x 6 matrix. Keeping in mind that $y = y(x_c, \dot{x}_c)$ equation (52) can be written

$$\frac{d}{dt} y^\dagger M^{-1} \frac{\partial y}{\partial \dot{x}_c} - y^\dagger M^{-1} \frac{\partial y}{\partial x_c} + \frac{\partial U(x_c)}{\partial x_c} = 0 \quad (54)$$

The derivatives $\frac{\partial y}{\partial x_c}$ and $\frac{\partial y}{\partial \dot{x}_c}$ are 6 x 6 matrices. Evaluating the time derivative gives

$$y^\dagger M^{-1} \frac{\partial y}{\partial \dot{x}_c} + y^\dagger \left(M^{-1} \frac{d}{dt} \frac{\partial y}{\partial \dot{x}_c} - M^{-1} \frac{\partial y}{\partial x_c} \right) + \frac{\partial U}{\partial x_c}(x_c) = 0 \quad (55)$$

The following definitions are useful:

$$A_{\ell k} \equiv \sum_{m=1}^6 M_{\ell m}^{-1} \frac{\partial y_m}{\partial \dot{x}_{ck}} \quad (56)$$

$$B_{\ell k} \equiv \sum_{m=1}^6 M_{\ell m}^{-1} \left(\frac{d}{dt} \frac{\partial y_m}{\partial \dot{x}_{ck}} - \frac{\partial y_m}{\partial x_{ck}} \right)$$

Equation (54) can be solved for \dot{y} . The result is

$$\dot{y}^\dagger = - \frac{\partial U(x)}{\partial x} A^{-1} - y^\dagger B A^{-1} \quad (57)$$

$\partial U / \partial x = \partial U / \partial x_c$ in agreement with (51) since the potential U does not depend on the momenta. From the transformation $y = y(x_c, \dot{x}_c)$ given by (37) and (44) one obtains for the matrix A

$$A = \left(\begin{array}{c|ccc} R^{-1}(\phi, \theta, \psi) & & & 0 \\ \hline & \sin \theta \sin \psi & \cos \psi & 0 \\ & \sin \theta \cos \psi & -\sin \psi & 0 \\ & \cos \theta & 0 & 1 \end{array} \right) \quad (58)$$

The inverse of this matrix is

$$A^{-1} = \left(\begin{array}{c|ccc} R(\phi, \theta, \psi) & & & 0 \\ \hline 0 & \frac{1}{\sin \theta} \sin \psi & \frac{1}{\sin \theta} \cos \psi & 0 \\ & \cos \psi & -\sin \psi & 0 \\ & -\text{ctg } \theta \sin \psi & -\text{ctg } \theta \cos \psi & 1 \end{array} \right) \quad (59)$$

We can write the matrix A and B in block form

$$A = \begin{pmatrix} R^{-1} & 0 \\ 0 & A' \end{pmatrix} \quad B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \quad (60)$$

$$BA^{-1} = \begin{pmatrix} B_1 R & B_2 A'^{-1} \\ B_3 R & B_4 A'^{-1} \end{pmatrix}$$

Comparison of (37)(44) with (58) gives

$$y = MA(x_c) \dot{x}_c \quad (61)$$

With (61) the matrix B can be expressed in terms of A.

$B = \frac{d}{dt} A - \frac{\partial}{\partial x_c} A \dot{x}_c$. The matrix $B_1 R$ is therefore equal to $(\frac{d}{dt} R^{-1})R = -\Omega$.

By direct calculation we find that also $B_4 A'^{-1}$ is equal to $-\Omega$. The matrix

B_3 vanishes. This leads to

$$BA^{-1} = - \begin{pmatrix} \Omega & \left(\frac{\partial}{\partial \alpha_j} \sum_{j,k} R_{jk}^{-1} \dot{q}_k A'_{il}^{-1} \right) \\ 0 & \Omega \end{pmatrix} \quad (62)$$

We define the differential operator D_x

$$D_x \equiv A^{-1} \frac{\partial}{\partial x} \quad (63)$$

According to (57) \dot{y} is

$$\dot{y} = -D_x U(x) + \begin{pmatrix} p'x\omega' \\ L'x\omega' \end{pmatrix} \cdot \quad (64)$$

We used the fact that the following contribution vanishes

$$\begin{aligned} \sum_{\ell, k} R_{i\ell}^{-1} \dot{q}_\ell \frac{\partial}{\partial \alpha_j} \left(R_{ik}^{-1} \dot{q}_k \right) &= \frac{1}{2} \sum_{\ell, k} \frac{\partial}{\partial \alpha_j} \left(R_{i\ell}^{-1} \dot{q}_\ell R_{ik}^{-1} \dot{q}_k \right) \\ &= \frac{1}{2} \frac{\partial}{\partial \alpha_j} || R^{-1} \dot{q} ||^2 = \frac{1}{2} \frac{\partial}{\partial \alpha_j} || \dot{q} ||^2 = 0 \end{aligned}$$

Equation (64) is Euler's equation of motion for a rigid body. The differential operator D_x is explicitly given by equations (111) and (112).

In the following it is more convenient to write the last term in equation (64) as a quadratic form in y

$$\begin{aligned} (\dot{y})_n &= -(D_x U(x))_n + \sum_{\ell, m} a_{\ell mn} y_\ell y_m \\ a_{\ell mn} &= \frac{1}{2} \left(C^{(n)} M^{-1} + M^{-1} C^{(n)\dagger} \right)_{\ell m} \end{aligned} \quad (65)$$

$$C^{(n)} = \begin{pmatrix} 0 & T_n \\ 0 & 0 \end{pmatrix}$$

$$C^{(n+3)} = \begin{pmatrix} 0 & 0 \\ 0 & T_n \end{pmatrix}$$

$n=1,2,3$

The tensor $a_{\ell mn}$ is defined such that $a_{\ell mn} = a_{m \ell n}$. With these definitions we obtain Liouville's equation (48) in the form we will use it in the following.

$$\frac{\partial}{\partial t} f(t, x, y) = \left\{ -yM^{-1} \cdot D_x + (D_x U(x)) \cdot \nabla - \sum_{\ell, m, n} a_{\ell mn} y_\ell y_m \nabla_n \right\} f(t, x, y) \quad (66)$$

The operator $\dot{x} \cdot \frac{\partial}{\partial x}$ in (48), (51) is equal to $yM^{-1} \cdot D_x$ since $y = M A(x) \dot{x}$ (61), (63). ∇ denotes the gradient with respect to y with components

$$\nabla_n \equiv \frac{\partial}{\partial y_n} .$$

Kramers-Liouville Equation

The motion of the particle is influenced by an external potential U and a "Brownian fluid", which is composed of molecules which exert fluctuating forces and torques

$$\tilde{h}(t) = (\tilde{F}(t), \tilde{N}(t)) \quad (67)$$

In absence of an external potential the equation of motion is

$$\dot{y} = - \int_{-\infty}^t ds \Gamma(t-s) y(s) + \tilde{h}(t) \quad (68)$$

For a derivation of the generalized Langevin equation (68) see [6]. The friction tensor $\Gamma(t)$ is proportional to the correlation of the fluctuating forces $\tilde{h}(t)$

$$\Gamma(t) = \frac{1}{kT} \langle \tilde{h}(0), \tilde{h}(t) \rangle \quad (69)$$

The symmetric tensor $\Gamma(t)$ is independent of the momenta y for heavy solute molecules. In the following we will use the "Markovian limit".

$$\dot{y} = -Cy + \tilde{h}(t) \quad C \equiv \int_0^{\infty} \Gamma(s) ds \quad (70)$$

The following discussion can be generalized simply by replacing the 6×6 matrix C with the corresponding expression in (68) in all equations.

In [6] equation (68) was derived from a linearized set of the equation of motion. Therefore one does not have to distinguish between the laboratory and the body fixed coordinate frames. The difference consists of quadratic terms $L' \times \omega'$ and $p' \times \omega'$. The idea is that over a short time of the order of the relaxation time both frames don't differ very much. After combining the stochastic equation (70) with Newton's equation, we can follow the orbit over an arbitrary long time and must therefore distinguish between both coordinate frames. The equation of motion containing the forces due to the fluid and the external forces is:

$$\dot{y} = -Cy + D_x U + \begin{pmatrix} p' x \omega' \\ L' x \omega' \end{pmatrix} + \tilde{h}(t) \quad (71)$$

In [12] and [13] the term $p' x \omega'$ is omitted. The generalization of Liouville's equation including stochastic forces can be obtained from (71)[1]. The result is the Kramers-Liouville equation

$$\begin{aligned} \frac{\partial}{\partial t} f &= (L + K)f \\ Lf &= -yM^{-1} \cdot D_x f + (D_x U) \cdot \nabla f - \sum_{\ell, m, n} a_{\ell mn} y_{\ell} y_m \nabla_n f \\ Kf &= \nabla \cdot C (M^{-1} y + kT \nabla) f \end{aligned} \quad (72)$$

The operator K is known as Kramers operator.

The Operator \tilde{L}

In the translational case it proved very useful to go to the "interaction picture".

$$f = e^{tK} \tilde{f}$$

$$\tilde{L}(t) \equiv e^{-tK} L e^{tK} = L + \sum_{n=1}^{\infty} [K, \cdot]^n L \left(\frac{(-t)^n}{n!} \right) \quad (73)$$

The operator L consists of three terms.

$$L = L_o + L_f + L_q$$

$$L_o = -y \cdot M^{-1} D_x$$

$$L_f = (D_x U) \cdot \nabla$$

$$L_q = - \sum_{lmn} a_{lmn} y_l y_m \nabla_n \quad (74)$$

The calculation of the operators \tilde{L}_o and \tilde{L}_f does not pose any difficulties. However, for \tilde{L}_q the situation is different since \tilde{L}_q contains quadratic terms in q. The commutators with K become more complicated.

All operators needed in (74) are contained in the algebra generated by $x_l, y_m, \nabla_n, \frac{\partial}{\partial x_i}$. The position and momenta are independent. From the definition (51) we obtain $[\nabla_n, x_l] = 0$ and $[\frac{\partial}{\partial x_i}, y_m] = 0$. The partial derivative $\frac{\partial}{\partial x_i}$ is evaluated with the momenta $y = (p', L')$ held constant. The differential operator $(D_x)_i$ (63) also commutes with y_m and ∇_n for all components i, m, n . The only nonvanishing commutator needed for the calculation of \tilde{L} is

$$[\nabla_n, y_m] = \delta_{nm} \quad n, m = 1, \dots, 6 \quad (75)$$

The operator $\tilde{L}_0(f)$ is given by the infinite sum $\tilde{L}_0(t) = L_0 + \sum_{n=1}^{\infty} [K, \cdot]^n L_0 \frac{(-t)^n}{n!}$. In order to simplify the notation we introduce the matrices \bar{C} and \tilde{C} and the operator \bar{D}_x

$$\bar{C} \equiv CM^{-1}, \quad \tilde{C} \equiv C kT, \quad \bar{D}_x \equiv -M^{-1} D_x \quad (76)$$

Kramers operator becomes

$$K = \nabla \cdot \bar{C} y + \nabla \cdot \tilde{C} \nabla. \quad (77)$$

The operator L_0 is $L_0 = y \cdot \bar{D}_x$. The first time dependent term in the expression for $\tilde{L}_0(t)$ is equal to $-t[K, L_0]$. This commutator is

$$\begin{aligned} [K, L_0] &= [\nabla \cdot \bar{C} y, y \cdot \bar{D}_x] + [\nabla \cdot \tilde{C} \nabla, y \cdot \bar{D}_x] \\ &= \sum_{n,l,m} \bar{C}_{nl} (\bar{D}_x)_m [\nabla_n y_l, y_m] \\ &+ \sum_{n,l,m} \tilde{C}_{nl} (\bar{D}_x)_m [\nabla_n \nabla_l, y_m]. \end{aligned} \quad (78)$$

The following identities hold for arbitrary operators A,B,C.

$$\begin{aligned} [A, BC] &= [A, B]C + B[A, C] \\ [AB, C] &= A[B, C] + [A, C]B \end{aligned} \quad (79)$$

With (79), (78) becomes

$$\begin{aligned}
 [K, L_0] = & \sum_{n, \ell, m} \bar{c}_{n\ell} (\bar{D}_x)_m \left\{ \nabla_n [y_\ell, y_m] + [\nabla_n, y_m] y_\ell \right\} \\
 & + \sum_{n, \ell, m} \check{c}_{n\ell} (\bar{D}_x)_m \left\{ \nabla_n [\nabla_\ell, y_m] + [\nabla_n, y_m] \nabla_\ell \right\}
 \end{aligned}$$

With (75) and using the fact that the matrix $\check{c} = C kT$ is symmetric [6], leads to

$$[K, L_0] = \bar{D}_x \cdot \bar{C} y + 2 \bar{D}_x \cdot \check{c} \nabla \tag{80}$$

For the higher order commutators one obtains

$$[K, \cdot]^n L_0 = \bar{D} \cdot \bar{C}^n y + 2 \sum_{m+\ell=n-1} \bar{D} \cdot \bar{C}^m \check{c} (-\bar{C}^\dagger)^\ell \nabla \tag{81}$$

This equation can be proved by induction on n . The calculation is similar to the calculation of $[K, L_0]$. We observe that the matrix $\bar{C}^m \check{c}$ is symmetric for all $m \geq 0$ since

$$\begin{aligned}
 \bar{C}^m \check{c} &= C M^{-1} C M^{-1} \dots C M^{-1} C kT \\
 &= (\bar{C}^m \check{c})^\dagger = \check{c} \bar{C}^{\dagger m}
 \end{aligned}$$

C and M are symmetric. Using this property the last term in (81) becomes

$2 \sum_{m+\ell=n-1} \bar{D}_x \cdot \bar{C}^m (-\bar{C})^\ell \check{c} \nabla$. The sum vanished for even n . For odd n it is equal to $2 \bar{D}_x \cdot \bar{C}^{n-1} \check{c} \nabla$.

$$[K, \cdot]^n L_0 = \begin{cases} \bar{D}_x \cdot \bar{C}^n y & , n \text{ even} \\ \bar{D}_x \cdot \bar{C}^n y + 2\bar{D}_x \cdot \bar{C}^{n-1} \hat{C} \nabla & , n \text{ odd} \end{cases} \quad (82)$$

The final result for the operator $\hat{L}_0(t)$ is

$$\begin{aligned} \hat{L}_0(t) &= \sum_{n=0}^{\infty} [K, \cdot]^n L_0 \left((-t)^n / n! \right) \\ &= \bar{D}_x \cdot e^{-t\bar{C}} y + \bar{D}_x \cdot (e^{t\bar{C}} - e^{-t\bar{C}}) \bar{C}^{-1} \hat{C} \nabla. \end{aligned} \quad (83)$$

and with the definitions of \bar{C} , \hat{C} and \bar{D}_x (76) one obtains

$$\begin{aligned} \hat{L}_0(t) &= -y \cdot M^{-1} E(-t) D_x \\ &+ kT \nabla \cdot [E(t) - E(-t)] D_x \end{aligned} \quad (84)$$

The matrix $E(t)$ is the exponential

$$E(t) \equiv e^{tCM^{-1}} \quad (85)$$

The corresponding expression used earlier for the translational motion

$-p/m \cdot \frac{\partial}{\partial q} e^{-\frac{\alpha}{m}t} + 2kT \sinh\left(\frac{\alpha}{m}t\right) \frac{\partial}{\partial p} \cdot \frac{\partial}{\partial q}$ is a special case of (84). It is remarkable that no higher than second order derivatives appear in $\hat{L}_0(t)$!

The calculation of the operator \hat{L}_f is similar. One obtains

$$\hat{L}_f(t) = \nabla \cdot E(t) [D_x U(x)] \quad (86)$$

In the final step we calculate the operator \hat{L}_q which is quadratic in the momenta y . This leads to major complications, but it turns out, that the operator $\hat{L}_q(t)$ contains no higher order derivatives than a third order derivative in the momenta q .

$\mathbb{1}$ is the 6 x 6 identity matrix. $W_{k\ell m}^{(n)}$ is symmetric in the first two indices $W_{k\ell m}^{(n)} = W_{\ell km}^{(n)}$ for all $n = 0, 1, 2, \dots$

Proof: All these relations follow directly from the definition of $X^{(n)}$, $Y^{(n)}$, $W^{(n)}$, $Z^{(n)}$ (88) and the definition of the commutator $[K \cdot]^n$ (16).

The following equations are true for arbitrary vectors $X^{(n)}$, $Y^{(n)}$, $W^{(n)}$, $Z^{(n)}$ with the only restriction that $W^{(n)}$ is symmetric in the first two indices.

$$W_{k\ell m}^{(n)} = W_{\ell km}^{(n)} \quad (91)$$

- (1) $[\nabla \cdot \bar{C}_y, X^{(n)*} \cdot \nabla] = (X^{(n)} \Omega)^* \cdot \nabla$
- (2) $[\nabla \cdot \bar{C}_y, W^{(n)} \cdot (y \otimes y \otimes \nabla)] = W^{(n)} \int \cdot (y \otimes y \otimes \nabla)$
- (3) $[\nabla \cdot \bar{C}_y, Y^{(n)} \cdot (y \otimes \nabla \otimes \nabla)] = Y^{(n)} T \cdot (y \otimes \nabla \otimes \nabla)$
- (4) $[\nabla \cdot \bar{C}_y, Z^{(n)} \cdot (\nabla \otimes \nabla \otimes \nabla)] = Z^{(n)} \phi \cdot (\nabla \otimes \nabla \otimes \nabla) \quad (92)$
- (5) $[\nabla \cdot \bar{C}_\nabla, W^{(n)} \cdot (y \otimes y \otimes \nabla)] = (W^{(n)} \Xi)^* \cdot \nabla + W^{(n)} \psi \cdot (y \otimes \nabla \otimes \nabla)$
- (6) $[\nabla \cdot \bar{C}_\nabla, Y^{(n)} \cdot (y \otimes \nabla \otimes \nabla)] = Y^{(n)} \Xi \cdot (\nabla \otimes \nabla \otimes \nabla)$
- (7) $[\nabla \cdot \bar{C}_\nabla, Z^{(n)} \cdot (\nabla \otimes \nabla \otimes \nabla)] = 0$

The proof of these equations is mostly straight forward. For instance the first equation is:

$$\begin{aligned} [\nabla \cdot \bar{C}_y, X^{(n)*} \cdot \nabla] &= \sum_{\alpha, \beta, \gamma} \bar{C}_{\alpha\beta} X_{\gamma}^{(n)*} [\nabla_{\alpha} y_{\beta}, \nabla_{\gamma}] \\ &= \sum_{\alpha, \beta, \gamma} \bar{C}_{\alpha\beta} X_{\gamma}^{(n)*} \nabla_{\alpha} (-\delta_{\alpha\beta}) = X^{(n)*} (-\bar{C}^{\dagger}) \cdot \nabla = (X^{(n)} \Omega)^* \cdot \nabla \end{aligned}$$

The 5th equation is different since there are two different type of terms:

$$\begin{aligned}
 & [\nabla \cdot \overset{\gamma}{C} \nabla, W^{(n)} \cdot (y \otimes y \otimes \nabla)] \\
 &= \sum_{\alpha, \beta, \gamma, \delta, \epsilon} \overset{\gamma}{C}_{\alpha\beta} W_{\gamma\delta\epsilon}^{(n)} \left\{ (\nabla_{\alpha} [\nabla_{\beta}, y_{\gamma}] + y_{\delta} \nabla_{\epsilon}) \right. \\
 &+ \left. [\nabla_{\alpha}, y_{\gamma}] \nabla_{\beta} y_{\delta} \nabla_{\epsilon} + y_{\gamma} \nabla_{\alpha} [\nabla_{\beta}, y_{\delta}] \nabla_{\epsilon} + y_{\gamma} [\nabla_{\alpha}, y_{\delta}] \nabla_{\beta} \nabla_{\epsilon} \right\}
 \end{aligned}$$

By assumption $W_{\gamma\delta\epsilon}^{(n)} = W_{\delta\gamma\epsilon}^{(n)}$ and $[y_{\alpha}, \nabla_{\beta}] = -\delta_{\alpha\beta}$. This gives the result $[\nabla \cdot \overset{\gamma}{C} \nabla, W^{(n)} \cdot (y \otimes y \otimes \nabla)] = (W^{(n)} \Xi)^* \cdot \nabla + W^{(n)} \Psi \cdot (y \otimes \nabla \otimes \nabla)$. The proof of the other equations is similar.

We define the vector valued function $W(t): \mathbb{R} \rightarrow \mathbb{R}^{216}$

$$W(t) \equiv \sum_{n=0}^{\infty} \left((-t)^n / n! \right) W^{(n)} \tag{93}$$

and similarly $X^*(t)$, $Y(t)$ and $Z(t)$. The recursion relations (89) for $W^{(n)}$, $X^{*(n)}$, $Y^{(n)}$ and $Z^{(n)}$ lead to the differential equations

$$W(0) = -a \quad , \quad X(0) = 0 \quad , \quad Y(0) = 0 \quad , \quad Z(0) = 0 \tag{94}$$

$$\begin{aligned}
 \dot{W}(t) &= -W(t) \sum \\
 \dot{X}(t) &= -X(t) \Omega - W(t) \Xi \\
 \dot{Y}(t) &= -Y(t) \Gamma - W(t) \Psi \\
 \dot{Z}(t) &= -Z(t) \phi - Y(t) \Xi
 \end{aligned}$$

These differential equations can be integrated and the results are

$$\begin{aligned}
 W(t) &= -a \exp(-t \Sigma) \\
 X(t) &= a \int_0^t ds \exp(-s \Sigma) \Xi \exp([s-t] \Omega) \\
 Y(t) &= a \int_0^t ds \exp(-s \Sigma) \Psi \exp([s-t] \Gamma) \\
 Z(t) &= - \int_0^t ds Y(s) \Xi \exp([s-t] \Phi)
 \end{aligned}
 \tag{95}$$

With these expressions the final result for the operator $\hat{L}(t)$ is with (84), (86), (88), (95):

$$\begin{aligned}
 \hat{L}(t) &= -y \cdot M^{-1} E(t) D_x \\
 &+ kT \nabla \cdot [E(t) - E(-t)] D_x \\
 &+ \nabla \cdot E(t) [D_x U(x)] \\
 &+ W(t) \cdot (y \otimes y \otimes \nabla) + X^*(t) \cdot \nabla \\
 &+ Y(t) \cdot (y \otimes \nabla \otimes \nabla) + Z(t) \cdot (\nabla \otimes \nabla \otimes \nabla)
 \end{aligned}
 \tag{96}$$

This is the Liouville operator in the interaction picture. The quadratic term L_q caused all the additional terms. Even if they are not explicitly known, we will be able to show that they do not contribute to the first and second cumulants.

First Cumulant

We calculate the cumulants under the assumption that initially the distribution in the momenta y is a Maxwell distribution

$$g(y) = \frac{1}{(2\pi kT)^3 (\det M)^{1/2}} e^{-\frac{y \cdot M^{-1} y}{2kT}} \tag{97}$$

The first cumulant is according to (25)

$$G^{(1)}(t) = \int d^6 y \tilde{L}(t) g(y) \quad (98)$$

We use expression (96) of $\tilde{L}(t)$ and integrate by parts. The contribution at the boundaries vanish. The remaining terms are integrals over odd functions in y_m , which vanish. The first cumulant is identical zero for all times $t \geq 0$

$$G^{(1)}(t) P(t, x) = 0 \quad (99)$$

Second Cumulant

The second cumulant gives the first nonvanishing contribution

$$G^{(2)}(t) = \int_0^t ds \int d^6 y \tilde{L}(t) \tilde{L}(s) g(y) \quad (100)$$

with (96)

$$\begin{aligned} G^{(2)}(t) = & - \int_0^t ds \int d^6 y y \cdot M^{-1} E(-t) D_x \\ & x \left[- y \cdot M^{-1} E(-s) D_x + kT \nabla \cdot \left\{ E(s) - E(-s) \right\} D_x \right. \\ & + \nabla \cdot E(s) [D_x U(x)] + W(s) \cdot (y \otimes y \otimes \nabla) \\ & \left. + X^*(s) \cdot \nabla + Y(s) \cdot (y \otimes \nabla \otimes \nabla) \right] g(y) \end{aligned} \quad (101)$$

The remaining terms of the product $\tilde{L}(t) \tilde{L}(s)$ vanish after integrating by parts. The only term left from the operator $\tilde{L}(t)$ is $-y \cdot M^{-1} E(-t) D_x$. Also the term $Z(s) \cdot (\nabla \otimes \nabla \otimes \nabla)$ vanished after integrating by parts three times.

At first we can show, that the contribution due to the terms $W(s) (y \otimes y \otimes \nabla)$, $X^*(s) \cdot \nabla$ and $Y(s) \cdot (y \otimes \nabla \otimes \nabla)$ cancel each other. We will show that the following integral vanishes for $k = 1, 2, \dots, 6$ and all times $s \geq 0$

$$J_k^{(s)} = \int d^6 y y_k [W(s) \cdot (y \otimes y \otimes \nabla) + X^*(s) \cdot \nabla + Y(s) \cdot (y \otimes \nabla \otimes \nabla)] g(y) \quad (102)$$

We recall that $\int d^6 y y_i y_j g(y) = M_{ij} kT$. Again integrating by parts (102) becomes

$$J_k(s) = - \sum_{n,m} (kT W_{nmk}(s) M_{nm} + X_{nnk}(s) - Y_{nnk}(s)) \quad (103)$$

The function $J_k(s)$ may be written as $J_k(s) = \sum_{n=0}^{\infty} J_k^{(n)} (-s)^n / n!$. For the constants $J_k^{(n)}$ one obtains according to (93)

$$J_k^{(\ell)} = - \sum_{n,m} \left(kT W_{nmk}^{(\ell)} M_{nm} + X_{nnk}^{(\ell)} - Y_{nnk}^{(\ell)} \right) \quad (104)$$

The recursion relations (89) allow us to define $J_k^{(\ell)}$ in terms of $J_k^{(\ell-1)}$

$$J_k^{(\ell)} = \sum_{\ell', k', m'} \left(kT W_{k'\ell'm'}^{(\ell-1)} M_{k'\ell'} \bar{C}_{m'k}^{\dagger} + X_{k'k'm'}^{(\ell-1)} \bar{C}_{m'k} + Y_{k'k'm'}^{(\ell-1)} \bar{C}_{m'k}^{\dagger} \right) \quad (105)$$

Comparing this expression with (103) shows

$$J_k^{(\ell)} = - \sum_{k'} J_k^{(\ell-1)} \bar{C}_{k'k}^{\dagger} = \left(J^{(\ell-1)} \bar{C}^{\dagger} \right)_k \quad (106)$$

The vector $J_k^{(0)}$ vanishes because $X^{(0)} = Y^{(0)} = 0$ and $\sum_{k'\ell'} W_{k'\ell'm}^{(0)} M_{k'\ell'} = - \sum_{k'\ell'} a_{k'\ell'm} M_{k'\ell'} = -\frac{1}{2} \text{Trace} (C^{(m)} + M^{-1} C^{(m)\dagger} M) = 0$ (65) (89). This shows that $J_k^{(\ell)} = 0$ for all ℓ and k . Therefore

$$J_k(s) = 0 \quad s \geq 0 \quad (107)$$

The integration of the remaining four terms in (101) is straight forward. One has to keep in mind that the matrix $M^{-1} E(t)$ is symmetric.

The final result is

$$\frac{d}{dt} P(t, \mathbf{x}) \approx G^{(2)}(t) P(t, \mathbf{x}) = D_{\mathbf{x}} \cdot \Lambda(t) \left[D_{\mathbf{x}} + \frac{1}{kT} (D_{\mathbf{x}} U(\mathbf{x})) \right] P(t, \mathbf{x}) \quad (108)$$

The time dependent diffusion matrix is

$$\Lambda(t) = kT C^{-1} \left(\mathbb{1} - e^{-tCM^{-1}} \right) \quad t \geq 0 \quad (109)$$

Equation (108) is the generalized Smoluchowski equation for coupled translational and rotational diffusion. Since we started with a Maxwell distribution at $t=0$, the diffusion matrix $\Lambda(t)$ is time dependent. Equation (108) includes as a special case the translational diffusion and the rotational diffusion discussed in [5]. The operator $D_{\mathbf{x}}$ depends on the orientation $\alpha = (\phi, \theta, \psi)$.

$$D_{\mathbf{x}} \equiv \begin{pmatrix} D_{q_1} \\ D_{\alpha} \end{pmatrix} \quad (110)$$

$$D_{q_1} = R^{\dagger}(\phi, \theta, \psi) \begin{pmatrix} \frac{\partial}{\partial q_1} \\ \frac{\partial}{\partial q_2} \\ \frac{\partial}{\partial q_3} \end{pmatrix} \quad (111)$$

$$D_{\alpha} = \begin{pmatrix} \cos \psi \frac{\partial}{\partial \theta} + \sin \psi \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} - \text{ctg } \theta \sin \psi \frac{\partial}{\partial \psi} \\ -\sin \psi \frac{\partial}{\partial \theta} + \cos \psi \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} - \text{ctg } \theta \cos \psi \frac{\partial}{\partial \psi} \\ \frac{\partial}{\partial \psi} \end{pmatrix} \quad (112)$$

The rotation $R(\phi, \theta, \psi)$ is defined in (33), (34). The expression for D_α follows from (59) and (63). Usually the friction tensor C is split into four 3×3 matrices.

$$C = \begin{pmatrix} C_{TT} & C_{TR} \\ C_{RT} & C_{RR} \end{pmatrix} \quad (113)$$

For axialsymmetric molecules it is easy to show that $C_{TR} = C_{RT} = 0$ [8]. In this case the diffusion equation is

$$\frac{\partial}{\partial t} P(t, q, \alpha) = \left\{ D_q \cdot \Lambda_T(D_q + \frac{1}{kT} [D_q U(q, \alpha)]) + D_\alpha \cdot \Lambda_R(D_\alpha + \frac{1}{kT} [D_\alpha U(q, \alpha)]) \right\} P(t, q, \alpha) \quad (114)$$

with

$$\Lambda_T = kT C_{TT}^{-1} \left(\mathbb{I} - e^{-t C_{TT}} \mathbb{I}^{-1} \right)$$

$$\Lambda_R = kT C_{RR}^{-1} \left(\mathbb{I} - e^{-t C_{RR}} \mathbb{I}^{-1} \right) \quad t \geq 0$$

The diffusion of translational and rotational degrees of freedom is still coupled even if the potential U vanishes, since D_q depends on α . In Section V we will solve (114) in two dimensions for $U(q, \alpha) = 0$.

In [12] and [13], different expressions for the operators corresponding to D_q and D_α , which are wrong in our opinion, are used. Instead of D_α the operator $J \equiv -i q \times \frac{\partial}{\partial q}$ was used. J is, up to a constant factor, the quantum mechanical angular momentum operator for a rotating point particle. Both operators D_α and J have the same commutator algebra since they are both infinitesimal generators of a representation of $SO(3)$. D_α and J correspond to two different representations; see (136). A connection between $J = -i q \times \frac{\partial}{\partial q}$ and the three Euler angles (ϕ, θ, ψ) also used in [12] and [13] is not obvious.

For axially symmetric particles one can factorize the angular dependence of $P(t, q, \phi, \theta, \psi)$ in ψ . The operator D_α^2 is in general not equal to $\Delta|_{r=1}$ the Laplace operator in spherical coordinates on the unit sphere. This is only true if we set $\frac{\partial}{\partial \psi} = 0$. If we consider only axial symmetric molecules and do not distinguish between two orientations which differ only by a rotation about the axis of symmetry, then we may use $D_\alpha^2|_\psi = \Delta|_{r=1}$; see (136). [12] obtained wrong results by setting $J^2 = \Delta$.

It is important to keep in mind that the operator D_q depends on the orientation. D_q is the gradient along the body fixed coordinate axis. If D_q is replaced by $D_q = \frac{\partial}{\partial q}$ one obtains wrong results [12], [13]. The coupling of translational and rotational diffusion of the two dimensional model discussed in Section V is a consequence of the α dependence of D_q , only.

These claims will be justified in detail in Section V.

IV. N Particle Diffusion

We consider N particles moving in a fluid interacting via arbitrary forces. In general the N particle density $P(t, x^{(1)}, x^{(2)}, \dots, x^{(N)})$ is not the product of the distributions $P(t, x^{(i)})$ where $x^{(i)}$ denotes the six coordinates of the i^{th} particle $x^{(i)} = (q^{(i)}, \alpha^{(i)})$. The N particles are correlated. The interaction energy is

$$U(x^{(1)}, x^{(2)}, \dots, x^{(N)}) \quad (114)$$

For an arbitrary observable $O(x^{(1)}, x^{(2)}, \dots, x^{(N)})$ depending on the position and orientation of the particles 1, ..., N the expectation value is defined

$$EO(t) \equiv \int d\mu_{\mathcal{X}} P(t, \mathcal{X}) O(\mathcal{X}) \quad (115)$$

with $\mathcal{X} \equiv (x^{(1)}, x^{(2)}, \dots, x^{(N)})$. The volume element $d\mu_{\mathcal{X}}$ is the product measure

$$d\mu_{\mathcal{X}} = \prod_{i=1}^N dq_1^{(i)} dq_2^{(i)} dq_3^{(i)} d\phi^{(i)} \sin \theta^{(i)} d\theta^{(i)} d\psi^{(i)} \quad (116)$$

The objective of this section is to derive the evolution equation for the N particle density $P(t, \mathcal{X})$ based on the Kramers-Liouville equation for the N particle motion. For the complete description of the N particle dynamics all positions $x^{(i)}$ and all momenta $y^{(i)}$ are required.

$$z^{(i)} \equiv (x^{(i)}, y^{(i)}) \quad (117)$$

$$z_{\mathcal{X}}(t) \equiv (z^{(1)}(t), z^{(2)}(t), \dots, z^{(N)}(t))$$

These variables are connected with the canonical variables $z_c(t)$ through the transformation (37) and (44) applied on every single coordinate $z^{(i)}$, $i=1, \dots, N$

$$z(t) = z(z_c(t)) \equiv \left(z^{(1)}(z_c^{(1)}(t)), \dots, z^{(N)}(z_c^{(N)}(t)) \right) \quad (118)$$

Liouville's equation holds for the density $f_c(t, z_c)$ since the determinant of the Jacobian matrix of the flux $z_c(t)$ is equal to one as a consequence of Hamilton's equation.

$$(\dot{x}_c^{(i)})_k = \frac{\partial H}{\partial (y_c^{(i)})_k}, \quad (\dot{y}_c^{(i)})_k = - \frac{\partial H}{\partial (x_c^{(i)})_k} \quad (119)$$

for $k = 1, 2, \dots, 6$ and $i = 1, 2, \dots, N$.

The Hamilton function is $H(x_c, y_c) = \frac{1}{2} \sum_{i=1}^N y^{(i)} \cdot M^{(i)-1} y^{(i)} + U(x^{(1)}, \dots, x^{(N)})$.

The matrix $M^{(i)}$ is the generalized inertia matrix (53) of the i^{th} particle.

Liouville's equation is

$$\frac{\partial}{\partial t} f_c(t, z_c) + \dot{z}_c \cdot \frac{\partial}{\partial z_c} f_c(t, z_c) = 0 \quad (120)$$

\dot{z}_c is determined by (119). The expectation value of an observable $O(z_c)$ is obtained by

$$\langle O \rangle(t) = \int d\mu_c f_c(t, z_c) O(z_c) \quad (121)$$

$d\mu_c$ is the volume element in the phase space $(S_{x_c} \times S_{y_c})^{xN}$.

$$d\mu_c = \prod_{i=1}^N \prod_{k=1}^{12} d(z_c^{(i)})_k \quad (122)$$

Instead of the canonical variables z_c we use again z . The transformations of the density f_c , the observable O and the measure $d\mu_c$ are

$$\begin{aligned}
 f(t, z) &\equiv f_c(t, z_c(z)) \\
 O(z) &\equiv O_{z_c}(z) \\
 d\mu &\equiv \left| \text{Det} \frac{\partial z_c}{\partial z} \right| dz \\
 &= \prod_{i=1}^N \sin \theta^{(i)} \prod_{k=1}^{12} dz_k^{(i)}
 \end{aligned}
 \tag{123}$$

The expectation value of the function $O(z)$

$$EO(t) = \int d\mu f(t, z) O(z)
 \tag{124}$$

agrees with the definition (121).

The Kramers-Liouville equation for the N particle problem has the form

$$\frac{\partial}{\partial t} f(t, z) = -\dot{z} \cdot \frac{\partial}{\partial z} f(t, z) + \sum_{i=1}^N K^{(i)} f(t, z)
 \tag{125}$$

$K^{(i)}$ is the Kramers operator acting on the i^{th} particle

$$K^{(i)} \equiv \nabla^{(i)} \cdot C^{(i)} [M^{(i)-1} y^{(i)} + kT \nabla^{(i)}]
 \tag{126}$$

$$\nabla^{(i)} \equiv \frac{\partial}{\partial y^{(i)}}$$

The Kramers operator is the direct sum of the individual operators $K^{(i)}$ acting on the i^{th} particle. The forces due to the fluid are completely random and not correlated at different positions. [6] The correlation matrix of all components

of all random forces and random torques, which is $6^N \times 6^N$ matrix, is the direct sum of the correlation matrices $C^{(i)}$. Therefore equation (126) is justified. With $L^{(i)}$, the Liouville operator acting on the i^{th} particle, the Kramers-Liouville equation (125) is the sum of N formally identical operators

$$\frac{\partial}{\partial t} f(t, z) = \sum_{i=1}^N (L^{(i)} + K^{(i)}) f(t, z)$$

$$L^{(i)} = -y^{(i)} \cdot M^{(i)-1} D_{x^{(i)}} + D_{x^{(i)}} U(x) \cdot \nabla^{(i)} \quad (127)$$

$$- a^{(i)} \cdot (y^{(i)} \otimes y^{(i)} \otimes \nabla^{(i)})$$

All operators $L^{(i)}$ are connected through the potential $U(x)$. Equation (127) contains the complete N body dynamics.

Since $[K^{(i)}, L^{(j)}] = 0$ for $i \neq j$ we have

$$e^{-t \sum_{i=1}^N K^{(i)}} \sum_{j=1}^N L^{(j)} e^{t \sum_{i=1}^N K^{(i)}} = \sum_{i=1}^N e^{-tK^{(i)}} L^{(i)} e^{tK^{(i)}} \quad (128)$$

$$= \sum_{i=1}^N \tilde{L}^{(i)}(t)$$

The operator $\tilde{L}^{(i)}(t)$ are given in equation (96) after replacing z by $z^{(i)}$, and M by $M^{(i)}$. The evolution equation for the density \tilde{f} defined by $f \equiv e^{tK} \tilde{f}$ is therefore

$$\frac{\partial}{\partial t} \tilde{f}(t, z) = \sum_{i=1}^N \tilde{L}^{(i)}(t) \tilde{f}(t, z) \quad (129)$$

Suppose the momentum distribution is Gaussian initially

$$g(\chi) = \prod_{i=1}^N g(y^{(i)})$$

$$g(y^{(i)}) = \frac{1}{(2\pi kT)^3 (\det M)^{1/2}} e^{-\frac{y^{(i)} \cdot M^{(i)-1} y^{(i)}}{2kT}} \quad (130)$$

As in the one particle case the first cumulant vanishes.

$$G^{(1)}(t) P(t, \chi) = \int \prod_{i=1}^N d^6 y^{(i)} \sum_{j=1}^N \dot{L}^{(j)}(t) \prod_{k=1}^N g(y^{(k)}) P(t, \chi) \quad (131)$$

$$= \sum_{j=1}^N \int d^6 y^{(j)} \dot{L}^{(j)}(t) g(y^{(j)}) P(t, \chi) = 0$$

The second cumulant is

$$G^{(2)}(t) P(t, \chi) = \int_0^t ds \int \prod_{i=1}^N d^6 y^{(i)} \sum_{\ell=1}^N \sum_{m=1}^N \dot{L}^{(\ell)}(t) \dot{L}^{(m)}(s) \prod_{j=1}^N g(y^{(j)}) P(t, \chi)$$

$$= \int_0^t ds \sum_{\ell=1}^N \int \prod_{i=1}^N d^6 y^{(i)} \dot{L}^{(\ell)}(t) \dot{L}^{(\ell)}(s) \prod_{j=1}^N g(y^{(j)}) P(t, \chi)$$

$$+ \int_0^t ds \sum_{\ell \neq m}^N \int \prod_{i=1}^N d^6 y^{(i)} \dot{L}^{(\ell)}(t) \dot{L}^{(m)}(s) \prod_{j=1}^N g(y^{(j)}) P(t, \chi)$$

$$= \left\{ \sum_{\ell=1}^N G^{(2)}(\ell)(t) + \sum_{\ell \neq m}^N \int_0^t ds G^{(1)}(\ell)(t) G^{(1)}(m)(s) \right\} P(t, \chi) \quad (132)$$

The second term vanishes because all first cumulants $G^{(1)}(\ell)$ $\ell=1 \dots N$ are zero.

The remaining term is the sum of the cumulants calculated for the one particle dynamics. The N particle diffusion equation is:

$$\frac{\partial}{\partial t} P(t, x^{(1)}, \dots, x^{(N)}) = \sum_{i=1}^N D_{x^{(i)}} \cdot \Lambda^{(i)}(t) \left[D_{x^{(i)}} + \frac{1}{kT} D_{x^{(i)}} U(x^{(1)}, \dots, x^{(N)}) \right] \\ \times P(t, x^{(1)}, \dots, x^{(N)}) \quad (133)$$

$$\Lambda^{(i)}(t) = kT C^{(i)-1} \left(1 - e^{-t C^{(i)} M^{(i)-1}} \right)$$

This is the generalization of the Smoluchowski equation for N interacting translating and rotating particles.

V. Correlations between the variables q and α .

We consider the one particle diffusion equation (114). In general the positions and orientations are correlated. The correlations are not only caused by the potential $U = U(x)$, $x = (q, \alpha)$ or by nonvanishing elements of the matrix $C_{TR} = C_{RT}^\dagger$. We will show that, if the positions q and the orientations α are uncorrelated at $t=t_0$ there are in general correlations for $t > t_0$ even if the potential vanishes and also $C_{TR}=0$.

Axially Symmetric Particles

As an illustration we consider axially symmetric particles. In this case one can show that $C_{TR} = 0$, [8]. If we identify the axis of symmetry with the e_3' axis the matrices C_{TT}^{-1} and C_{RR}^{-1} are diagonal.

$$C_{TT}^{-1} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a_3 \end{pmatrix} \quad C_{RR}^{-1} = \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b_3 \end{pmatrix} \quad (134)$$

We assume that we know the distribution at time $t = t_0$, where t_0 is large compared with the translational and rotational relaxation time of the momenta.

$$t_0 \gg \tau \quad || C_{TT}^{-1} || \quad \text{and} \quad t_0 \gg \tau \quad || C_{RR}^{-1} I || \quad (135)$$

$$\frac{\partial}{\partial t} P(t, q, \alpha) = kT \left[a D_q^2 + (a_3 - a) (D_q)_3^2 + b D_\alpha^2 + (b_3 - b) (D_\alpha)_3^2 \right] P(t, q, \alpha)$$

for $t > t_0$

This equation is a special case of (114) where we used expression (134) for the friction tensor. We also used $\Lambda(t) \approx kT C^{-1}$ for $t \geq t_0$.

The differential operators $(D_q,)^2$, $(D_q,)_3^2$, D_α^2 and $(D_\alpha)_3^2$ are given by equation (111) and (112).

$$\begin{aligned}
 (D_q,)^2 &= \Delta_q \\
 (D_q,)_3^2 &= \frac{\partial}{\partial q} \cdot B(\alpha) \frac{\partial}{\partial q} \quad ; \quad B_{ij}(\alpha) \equiv R_{3i}(\alpha) R_{3j}(\alpha) \\
 (D_\alpha)^2 &= \frac{\partial^2}{\partial^2 \theta} + \frac{1}{\sin^2 \theta} \left(\frac{\partial^2}{\partial^2 \phi} + \frac{\partial^2}{\partial^2 \psi} \right) \\
 &\quad + 2 \frac{\cos \theta}{\sin^2 \theta} \frac{\partial}{\partial \phi} \frac{\partial}{\partial \psi} + \text{ctg } \theta \frac{\partial}{\partial \theta} \\
 (D_\alpha)_3^2 &= \frac{\partial^2}{\partial^2 \psi}
 \end{aligned} \tag{136}$$

Δ_q is the Laplace operator in Cartesian coordinates.

We define the new density $P(t, q, \phi, \theta)$

$$P(t, q, \phi, \theta) = \int d\psi P(t, q, \phi, \theta, \psi). \tag{137}$$

Integrating equation (135) on both sides with respect to ψ leads to

$$\begin{aligned}
 \frac{\partial}{\partial t} P(t, q, \phi, \theta) &= kT \left[a \Delta_q + (a_3 - a) \frac{\partial}{\partial q} \cdot B(\phi, \theta) \frac{\partial}{\partial q} \right. \\
 &\quad \left. + b \Delta \Big|_{r=1} \right] P(t, q, \phi, \theta)
 \end{aligned} \tag{138}$$

The matrix $B(\alpha)$ defined in equation (136) does not depend on ψ .

$$B(\phi, \theta) = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} \otimes \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix} \tag{139}$$

The contributions of equation (135) which contain a derivative with respect to ψ vanish after integrating by parts. Therefore the operator D_α^2 reduces to $\Delta|_{r=1}$, the Laplace operator in spherical coordinates on the unit sphere.

$$D_\alpha^2|_\psi = \Delta|_{r=1} \quad (140)$$

We assume that the initial condition factorizes. For $t > t_0$ the solution of (138) has the form

$$P(t_0, q, \phi, \theta) = P_{OT}(q) P_{OR}(\phi, \theta)$$

$$P(t, q, \phi, \theta) = P_T(t, q, [P_R]) P_R(t, \phi, \theta) \quad (141)$$

$$t > t_0$$

The function $P_T(t)$ is also a functional of the distribution $P_R(t)$. $P_T(t)$ and $P_R(t)$ are probability densities, $\int d^3q P(t, q, [P_R]) = 1$ and $\int d\phi d\theta \sin \theta P_R(t, \phi, \theta) = 1$ for all times $t > t_0$. The boundary conditions are: $P_T(t, q, [P_R]) = 0$ if $q_i = \infty$ for some $i=1,2,3$. Substituting (141) into equation (138) and integrating with respect to ϕ and θ (using the weight $\sin \theta$) leads to equation (142). Similarly one obtains (143) by integrating with respect to q .

$$\begin{aligned} \frac{\partial}{\partial t} P_T(t, q, [P_R]) = kT \left\{ a\Delta_q \right. \\ \left. + (a_3 - a) \sum_{i,j} \int d\phi \sin \theta d\theta B_{ij}(\phi, \theta) P_R(t, \phi, \theta) \frac{\partial}{\partial q_i} \frac{\partial}{\partial q_j} \right\} \\ \times P_T(t, q, [P_R]) \end{aligned} \quad (142)$$

$$\frac{\partial}{\partial t} P_R(t, \phi, \theta) = kT b \Delta|_{r=1} P_R(t, \phi, \theta) \quad (143)$$

$$\text{for } t \geq t_0$$

The second equation describes the "Brownian motion on the unit sphere".

The eigenfunction of $\Delta|_{r=1}$ are the spherical harmonics $Y_{lm}(\theta, \phi)$. Substituting a solution $P_R(t, \phi, \theta)$ of (143) into equation (142) one obtains an expression, which is formally a diffusion equation with time dependent diffusion coefficients. The off diagonal elements of the diffusion matrix vanish if the distribution $P_R(t, \phi, \theta)$ is uniform.

Similarly, one can show that for arbitrary molecules with $C_{TR} = 0$ a solution of the form (141) (including ψ) exists, if the positions and orientations are uncorrelated at time $t = t_0$ and if $U = 0$.

Diffusion in Two Dimensions

In two dimensions the diffusion equation without external potential can be solved for arbitrary initial conditions. Equation (108) reduces to

$$\frac{\partial}{\partial t} P(t, q_1, q_2, \phi) = A P(t, q_1, q_2, \phi) \quad (144)$$

$$A = \begin{pmatrix} \frac{\partial}{\partial q_1} \\ \frac{\partial}{\partial q_2} \end{pmatrix} \cdot \Lambda(\phi) \begin{pmatrix} \frac{\partial}{\partial q_1} \\ \frac{\partial}{\partial q_2} \end{pmatrix} + kT \gamma \frac{\partial^2}{\partial \phi^2} \quad t \geq t_0$$

$$\Lambda(\phi) = kT \begin{pmatrix} \alpha \cos^2 \phi + \beta \sin^2 \phi & (\beta - \alpha) \sin \phi \cos \phi \\ (\beta - \alpha) \sin \phi \cos \phi & \alpha \sin^2 \phi + \beta \cos^2 \phi \end{pmatrix} \quad (145)$$

$kT\alpha$, $kT\beta$, $kT\gamma$ are the diffusion constants corresponding to the degrees of freedom q_1 , q_2 and ϕ . We assume that $\alpha > \beta$. We use the following identities to simplify the matrix $\Lambda(\phi)$:

$$\begin{aligned} \alpha \cos^2 \phi + \beta \sin^2 \phi &= \frac{\alpha + \beta}{2} + \frac{\alpha - \beta}{2} \cos 2\phi \\ \alpha \sin^2 \phi + \beta \cos^2 \phi &= \frac{\alpha + \beta}{2} - \frac{\alpha - \beta}{2} \cos 2\phi \\ 2 \sin \phi \cos \phi &= \sin 2\phi \end{aligned} \quad (146)$$

$$\Lambda(\phi) = kT \delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + kT \epsilon \begin{pmatrix} \cos 2\phi & -\sin 2\phi \\ -\sin 2\phi & -\cos 2\phi \end{pmatrix} \quad (147)$$

δ is the average translational diffusion constant and ϵ is a measure for the asymmetry of the particle.

$$\delta \equiv \frac{\alpha + \beta}{2} \quad \epsilon \equiv \frac{\alpha - \beta}{2} \quad (148)$$

Without solving (144) explicitly it is already possible to make some statements about the lowest moments of q_1 , q_2 and ϕ . One obtains the following differential equations for the expectation values

$$\langle \dots \rangle_t = \int dq_1 dq_2 d\phi \dots P(t, q_1, q_2, \phi):$$

$$\begin{aligned}
 \frac{d}{dt} \langle q_1 \rangle_t &= 0 \\
 \frac{d}{dt} \langle q_1^2 \rangle_t &= 2kT \delta + 2kT \epsilon \langle \cos 2\phi \rangle_t \\
 \frac{d}{dt} \langle q_1 q_2 \rangle_t &= -2kT \epsilon \langle \sin 2\phi \rangle_t \\
 \frac{d}{dt} \langle \cos 2\phi \rangle_t &= -4\gamma kT \langle \cos 2\phi \rangle_t \\
 \frac{d}{dt} \langle \sin 2\phi \rangle_t &= -4\gamma kT \langle \sin 2\phi \rangle_t
 \end{aligned}
 \tag{149}$$

This leads to:

$$\begin{aligned}
 \langle \cos 2\phi \rangle_t &= e^{-4\gamma kTt} \langle \cos 2\phi \rangle_{t_0} \\
 \langle q_1^2 \rangle_t &= 2kT \delta t + \frac{\epsilon}{2\gamma} (1 - e^{-4\gamma kTt}) \langle \cos 2\phi \rangle_{t_0} + \langle q_1^2 \rangle_{t_0} \\
 \langle q_1 q_2 \rangle_t &= \frac{-\epsilon}{2\gamma} (1 - e^{-4\gamma kTt}) \langle \sin 2\phi \rangle_{t_0} + \langle q_1 q_2 \rangle_{t_0}
 \end{aligned}
 \tag{150}$$

The calculation of arbitrary expectation values $\langle O \rangle_t$, $O = O(q_1, q_2, \phi)$ can be reduced to the problem of finding the eigenvectors and eigenvalues of the diffusion operator A in equation (144).

$$(A - \lambda_{k_1 k_2 \ell}) \psi_{k_1 k_2 \ell} = 0
 \tag{151}$$

For the symmetric case $\alpha = \beta$ the solutions of (149) are

$$\begin{aligned}
 \psi_{k_1 k_2 \ell}^s(q_1, q_2, \phi) &= \frac{1}{\sqrt{\pi}L} e^{ik_1 q_1} e^{ik_2 q_2} \sin(\ell\phi) \\
 \psi_{k_1 k_2 \ell}^c(q_1, q_2, \phi) &= \frac{1}{\sqrt{\pi}L} e^{ik_1 q_1} e^{ik_2 q_2} \cos(\ell\phi)
 \end{aligned}
 \tag{152}$$

We choose a box of length L and assume periodic boundary conditions

$$\psi(0, q_2, \phi) = \psi(L, q_2, \phi) , \quad \psi(q_1, 0, \phi) = \psi(q_1, L, \phi) \quad (153)$$

$$\psi(q_1, q_2, \phi) = \psi(q_1, q_2, \phi + 2\pi)$$

The possible values for k_1 , k_2 and ℓ are

$$k_1 = \pm \frac{2n\pi}{L} , \quad k_2 = \pm \frac{2m\pi}{L} \quad n, m \in \mathbb{N} \quad (154)$$

$$\ell = 0, 1, 2, \dots$$

In the general case $\alpha > \beta$ we make the ansatz that the eigenfunctions can be written

$$\psi_{k_1 k_2 \ell}(q_1, q_2, \phi) = \frac{1}{\sqrt{L}} e^{ik_1 q_1} e^{ik_2 q_2} \varepsilon_{k_1 k_2 \ell}(\phi) \quad (155)$$

One obtains the following differential equation for the unknown function

$$\varepsilon_{k_1 k_2 \ell}(\phi) \quad (144), \quad (147), \quad (151)$$

$$\left[-\delta(k_1^2 + k_2^2) - \varepsilon(k_1^2 - k_2^2) \cos 2\phi \right. \\ \left. + 2 \varepsilon k_1 k_2 \sin 2\phi + \gamma \frac{\partial^2}{\partial \phi^2} - \frac{\lambda k_1 k_2 \ell}{kT} \right] \varepsilon_{k_1 k_2 \ell}(\phi) = 0 \quad (156)$$

We define the complex wave number k' .

$$k' \equiv k_1 + ik_2 \\ \psi \equiv \arctan \left(\frac{k_2}{k_1} \right) \quad (157)$$

k' can be written $k' = |k'| e^{i\psi}$. Equation (156) becomes:

$$\left\{ -\left(\delta |k'|^2 + \frac{\varepsilon}{2} |k'|^2 e^{2i\phi+2i\psi} + \frac{\varepsilon}{2} |k'|^2 e^{-2i\phi-2i\psi} \right) + \gamma \frac{\partial^2}{\partial \phi^2} - \frac{\lambda_{k_1 k_2 \ell}}{kT} \right\} g_{k_1 k_2 \ell}(\phi) = 0 \quad (158)$$

The exponentials can be combined to $\cos(2[\phi+\psi])$. Equation (158) is equivalent to Mathieu's equation [9].

$$\frac{d}{dz} y_{\ell}(z) + (a_{\ell}(r) - 2r \cos 2z) y_{\ell}(z) = 0 \quad (159)$$

$$r = \frac{(k_1^2 + k_2^2)(\alpha - \beta)}{4\gamma}$$

$$z = \phi + \arctan \left(\frac{k_2}{k_1} \right) \quad (160)$$

$$\lambda_{k_1 k_2 \ell} = -kT \left\{ \gamma a_{\ell}(r) + \frac{\alpha + \beta}{2} \left(k_1^2 + k_2^2 \right) \right\}$$

$$g_{k_1 k_2 \ell}(\phi) = y_{\ell} \left(\phi + \arctan \left(\frac{k_2}{k_1} \right) \right)$$

The eigenvalues $a_{\ell}(r)$ of Mathieu's equation are negative for certain values of r and ℓ [9], but the eigenvalues $\lambda_{k_1 k_2 \ell}$ are always less or equal to zero for all k_1, k_2 and ℓ .

Equation (159) has a complete set of orthogonal solutions $ce_{\ell}(r, z)$ and $se_{\ell}(r, z)$ with the corresponding eigenvalues denoted by $a_{\ell}(r)$ and $b_{\ell}(r)$ [9].

The eigenfunctions of (151) are

$$\begin{aligned} \psi'_{k_1 k_2 \ell}(q_1, q_2, \phi) &= \frac{1}{\sqrt{\pi} L} e^{ik_1 q_1} e^{ik_2 q_2} \operatorname{se}_\ell \left(r, \phi + \arctan \left(\frac{k_2}{k_1} \right) \right) \\ \psi_{k_1 k_2 \ell}(q_1, q_2, \phi) &= \frac{1}{\sqrt{\pi} L} e^{ik_1 q_1} e^{ik_2 q_2} \operatorname{ce}_\ell \left(r, \phi + \arctan \left(\frac{k_2}{k_1} \right) \right). \end{aligned} \quad (161)$$

since $\left\{ \psi'_{k_1 k_2 \ell}, \psi_{k_1 k_2 \ell} \right\}$ is a complete set of orthogonal eigenfunctions of the diffusion operator (151), the expectation value $\langle O \rangle_t$ can be found by

$$\begin{aligned} \langle O \rangle_t &= \int dq_1 dq_2 d\phi P(t, q_1, q_2, \phi) O(q_1, q_2, \phi) \\ &= \sum_{k_1 k_2 \ell} e^{\lambda_{k_1 k_2 \ell} t} P_{k_1 k_2 \ell} O_{k_1 k_2 \ell} \\ &\quad + \sum_{k_1 k_2 \ell} e^{\lambda'_{k_1 k_2 \ell} t} P'_{k_1 k_2 \ell} O'_{k_1 k_2 \ell} \end{aligned} \quad (162)$$

The coefficients $O_{k_1 k_2 \ell}$, $O'_{k_1 k_2 \ell}$, $P_{k_1 k_2 \ell}$, $P'_{k_1 k_2 \ell}$ are obtained from $O(q_1, q_2, \phi)$ and the initial distribution $P(t_0, q_1, q_2, \phi)$.

$$\begin{aligned} O_{k_1 k_2 \ell} &= \int dq_1 dq_2 d\phi \psi_{k_1 k_2 \ell}^*(q_1, q_2, \phi) O(q_1, q_2, \phi) \\ &= (\psi_{k_1 k_2 \ell}, O) \\ O'_{k_1 k_2 \ell} &= (\psi'_{k_1 k_2 \ell}, O) \\ P_{k_1 k_2 \ell} &= (P(t_0), \psi_{k_1 k_2 \ell}) \\ P'_{k_1 k_2 \ell} &= (P(t_0), \psi'_{k_1 k_2 \ell}) \end{aligned} \quad (163)$$

As an illustration we consider the following two observables:

$$O^S(q_1, \phi) \equiv \sin(k_1 q_1) se_1(r, \phi) \tag{164}$$

$$O^C(q_1, \phi) \equiv \sin(k_1 q_1) ce_1(r, \phi)$$

with $k_1 = 2\pi/L$ and $r = \frac{\pi^2(\alpha - \beta)}{\gamma L^2}$

We assume that the asymmetry is small. In this case $r \ll 1$ and the Mathieu functions se_1 and ce_1 are approximately

$$ce_1(r, \phi) \approx \cos(\phi) - \frac{r}{8} \cos(3\phi) \tag{165}$$

$$se_1(r, \phi) \approx \sin(\phi) - \frac{r}{8} \sin(3\phi)$$

The corresponding eigenvalues are

$$\begin{aligned} a_1(r) &\approx 1 + r \\ b_1(r) &\approx 1 - r \end{aligned} \tag{166}$$

The eigenvalues $\lambda_{\pm k_1 01}$ and $\lambda'_{\pm k_1 01}$ are

$$\begin{aligned} \lambda_{\pm k_1 01} &\approx -kT \left(\gamma + \frac{\pi^2}{L^2} \left(3\alpha + \beta \right) \right) \\ \lambda'_{\pm k_1 01} &\approx -kT \left(\gamma + \frac{\pi^2}{L^2} \left(\alpha + 3\beta \right) \right) \end{aligned} \tag{167}$$

and for the expectation values of O^S and O^C one obtains

$$\begin{aligned} \langle O^C \rangle_t &\approx c e^{-kT \left(\gamma + \frac{\pi^2}{L^2} \left(3\alpha + \beta \right) \right) t} \\ \langle O^S \rangle_t &\approx c' e^{-kT \left(\gamma + \frac{\pi^2}{L^2} \left(\alpha + 3\beta \right) \right) t} \end{aligned} \tag{168}$$

The constants c and c' can be written $c = (O^C, P(t))$ and $c' = (O^S, P(t))$.

The state O^c decays faster since we assumed $\alpha > \beta$. α corresponds to the diffusion along the e_1' axis of the molecule. In the state O^c the molecule axis e_1' is mainly parallel to the e_1 direction of the laboratory frame; in the state O^s e_1' is mainly parallel to the e_2 axis. The average speed of the molecules in state O^c is bigger in the direction e_1 ; e_1 is also the direction of the spatial inhomogeneity. Therefore O^c decays faster than O^s . This example is typical for the type of coupling of q_1, q_2 and ϕ , which occurs in the translational and rotational diffusion if the potential U vanishes and also $C_{TR} = 0$.

VI. Concluding Remarks

We have shown that a "contraction of the description" is achieved when a Kramers-Liouville process is averaged with respect to its momenta variables. The second cumulant of an ordered time evolution cumulant expansion yields the generalized Smoluchowski equation as the contracted description. We have examined the details of the dynamical operator algebra generated by the contraction procedure for translational and rotational degrees of freedom, and for as many as N distinct particles.

A more thorough description of the higher order cumulants, shown to be small here, will appear in a forthcoming paper.

Acknowledgement

An NSF Grant for theoretical physics has supported this research.

REFERENCES

- [1] R. F. Fox, Phys. Rep. 48, 179 (1978).
- [2] R. F. Fox, J. Math. Phys. 20, 2467 (1979).
- [3] R. F. Fox, J. Math. Phys. 17, 1148 (1976).
- [4] K. Yosida, Functional Analysis, (Springer Verlag 1974).
- [5] L. D. Favro, Phys. Rev. 119, 53 (1966).
- [6] E. H. Hauge, A. Martin-Löf, J. Stat. Phys. 7, 259 (1973).
- [7] D. W. Condiff, J. S. Dahler, J. Chem. Phys. 44, 3988 (1966).
- [8] J. Happel, H. Brenner, Low Reynolds Number Hydrodynamics, (Prentiss Hall Inc. 1965).
- [9] I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals and Series and Products, (Academic P., New York, 1965). Section 8.6.
- [10] R. Jost, Quantenmechanik I (Verein der Mathematiker und Physiker, ETH Zürich 1969)
- [11] V. I. Arnold Mathematical Methods of Classical Mechanics, (Springer Verlag. 1978)
- [12] G. T. Evans, J. Chem. Phys. 65, 3030 (1978).
- [13] L.-P. Hwang, J. H. Freed, J. Chem. Phys. 63, 118 (1975).
- [14] J. L. Davis, R. F. Fox, J. Stat. Phys. 22, 627 (1980).
- [15] S. Chandrasekhar, Rev. Mod. Phys. 15, 1, (1943).
- [16] H. Goldstein, Classical Mechanics, (Addison-Wesley, Reading, Massa., 1959) chapters 4 and 5.

The “excess entropy” around nonequilibrium steady states, $(\delta^2 S)_{ss}$, is not a Liapunov function

RONALD FORREST FOX

School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332

Communicated by Kenneth M. Case, April 14, 1980

ABSTRACT In a response to my recent paper [Fox, R. F. (1979) *Proc. Natl. Acad. Sci. USA* 76, 2114–2117], Nicolis and Prigogine [Nicolis, G. & Prigogine, I. (1979) *Proc. Natl. Acad. Sci. USA* 76, 6060–6061] reasserted that the “excess entropy” around nonequilibrium steady states, $(\delta^2 S)_{ss}$, is a Liapunov function. A simple, explicit counterexample which invalidates this claim is presented. The existence of an alternative theory possessing a proper Liapunov function for steady states is reviewed.

One way to describe the stability of nonequilibrium steady states is to construct a Liapunov function, $L(t)$, which satisfies the stability inequalities

$$\begin{aligned} L(t) &\leq 0 \\ \frac{dL(t)}{dt} &\geq 0 \end{aligned} \quad [1]$$

if and only if the steady state is stable. In the linear regime near full equilibrium, the second-order excess entropy, $(\delta^2 S)_{eq}$, provides such a function, as was discussed recently by Fox (1) and as has been repeatedly observed earlier by others, especially Prigogine (2). In addition, $(\delta^2 S)_{eq}$ near full equilibrium provides the Einstein–Boltzmann–Planck formula for near-equilibrium fluctuations, a fact that enabled Onsager to build his celebrated theory of near-equilibrium thermodynamics (3). These two properties of $(\delta^2 S)_{eq}$ are intimately connected by the fluctuation–dissipation theorem for near-equilibrium thermodynamics; and this fact was one of the main points of the recent paper by Fox (1). On purely mathematical grounds, many Liapunov functions could be constructed for near-equilibrium thermodynamics; but the particular choice of $(\delta^2 S)_{eq}$ is of special interest because of its additional physical significance, which it derives from its additional role in determining the near-equilibrium fluctuations.

In the case of nonequilibrium steady states, the question has been raised whether there exists a similar quantity. Prigogine has repeatedly suggested that the excess entropy around the steady state, $(\delta^2 S)_{ss}$, possesses the Liapunov property, although he has also shown that it does not provide the basis for a formula for the nonequilibrium fluctuations around the steady state. A theory for nonequilibrium steady-state fluctuations developed by Keizer (4, 5) has established that the covariance of the steady-state fluctuations can be used to construct a Liapunov function for steady states. This quantity parallels the behavior of $(\delta^2 S)_{eq}$ in that it provides both the steady-state fluctuations and a Liapunov criterion. Fox (1) reviewed the fact that it is the fluctuation–dissipation relation for steady states that is responsible for this dual role, just as is the case for $(\delta^2 S)_{eq}$ near full equilibrium. Of the many mathematically possible Liapunov functions for steady states, the choice of a Liapunov function based upon the covariance of steady-state fluctuations is espe-

cially attractive because of its additional physical significance. Indeed, Keizer (5, 6) has shown that it may be used to build a thermodynamics of steady states which parallels the thermodynamics at equilibrium in every detail.

Contrary to the claim of Nicolis and Prigogine (7), the problem of finding Liapunov functions at steady states is not merely a matter of choosing between $(\delta^2 S)_{ss}$ and the covariance of the steady-state fluctuations. The reason this is not the issue is that $(\delta^2 S)_{ss}$ is not always in fact a Liapunov function at steady states. So that there will be no doubt regarding the existence of the opposite claim—that is, that $(\delta^2 S)_{ss}$ is a Liapunov function for steady states—I will quote solely from the recent paper of Nicolis and Prigogine (7), although other sources for this claim are referenced in an earlier paper (1). The italics in the quotations are mine and the references are those in the paper of Nicolis and Prigogine (7). The first paragraph of ref. 7 states: “The central quantity in this theory is $(\delta^2 S)_{ss}$, the second-order excess entropy around the steady state, which is used as Lyapunov function to derive linear stability criteria.” The fifth paragraph of ref. 7 states: “In the light of these definitions of stability and of Lyapunov functions (7), $(\delta^2 S)_{ss}$ keeps its significance entirely. As we pointed out (4, 5), its usefulness is 2-fold: first, it has a macroscopic meaning below, at, and across bifurcation points, *independently* of the fine—and often complex—details of the behavior of the fluctuations; and second, it enjoys *universality*, as it can be applied to a wide class of systems including those subject to spatially inhomogeneous disturbances, surface effects, and so forth.” In paragraph six: “Naturally, the above summarized properties of $(\delta^2 S)_{ss}$ do not imply that there is an argument against the use of *other* Lyapunov functions.” In paragraph seven they conclude: “The covariance of the fluctuating thermodynamic variables advocated by Fox (3) provides an example of just such *another* Lyapunov function”

All of the preceding would carry weight if it were not for the fact, to be exhibited below, that $(\delta^2 S)_{ss}$ is *not* always a Liapunov function at steady states. That it is not one was observed by Keizer and Fox (8) some time ago. Evidently, we did not make this point sufficiently clear to bring an end to the debate which continues to this day (1, 7). In the following, an explicit counterexample is used to show explicitly that *along the solutions of the rate equations*

$$\begin{aligned} (\delta^2 S)_{ss} &\leq 0 \\ \frac{d(\delta^2 S)_{ss}}{dt} &< 0 \end{aligned} \quad [2]$$

for a steady state that is asymptotically stable. The stability will be shown on the basis of other stability criteria. With this example, there can no longer be any question regarding the universality enjoyed by $(\delta^2 S)_{ss}$ as a Liapunov function.

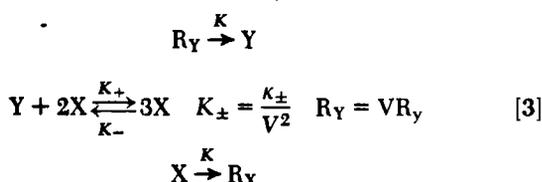
The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U. S. C. §1734 solely to indicate this fact.

COUNTEREXAMPLE

The purpose of this counterexample is to exhibit a simple system in which there is a stable steady state, but for which $(\delta^2 S)_{ss}$ fails as a Liapunov function, along the trajectory to the steady state of the solutions to the rate equations. It will be shown that, even though the approach to the steady state is completely stable, all along the trajectory of the solutions $d(\delta^2 S)_{ss}/dt < 0$ repeatedly, with a regular frequency. It is also true that $d(\delta^2 S)_{ss}/dt < 0$ for trajectories already arbitrarily close to the steady state. This failure of $(\delta^2 S)_{ss}$ as a Liapunov function for steady states is not shared by the choice of a *bona fide* Liapunov function in Keizer's theory (4, 5, 9).

The model describes the reaction dynamics of a driven, homogeneous-phase, chemical reaction which involves a modified Schlogl mechanism. Such a mechanism was the prominent feature in another model, already discussed by Keizer and Fox (8). We already showed then that $(\delta^2 S)_{ss}$ could not be used as a Liapunov function.

The reaction is



The rate equations for this reaction which is driven by the constant input K R_Y is given below in concentration variables: $x = \text{X}/V$ and $y = \text{Y}/V$.

$$\begin{aligned} \dot{x} &= \kappa_+ x^2 y - \kappa_- x^3 - \text{K}x \\ \dot{y} &= \text{K R}_Y - \kappa_+ x^2 y + \kappa_- x^3. \end{aligned} \quad [4]$$

The steady state values of x and y , for fixed K R_Y , are

$$\begin{aligned} x_{ss} &= \text{R}_Y \\ y_{ss} &= \alpha \text{R}_Y + \frac{\text{K}}{\kappa_+ \text{R}_Y} \quad \alpha \equiv \frac{\kappa_-}{\kappa_+}. \end{aligned} \quad [5]$$

The local equilibrium entropy per unit volume is the same as in equation 39 of ref. 1:

$$S(t) = S_0 - \text{K}_B (x \ln Vx + y \ln Vy). \quad [6]$$

The resulting second-order excess entropy is

$$\begin{aligned} \frac{1}{2} \delta^2 S &= -\frac{1}{2} \text{K}_B (\delta x, \delta y) \begin{pmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{pmatrix} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \\ &= -\frac{1}{2} \text{K}_B \left(\frac{(\delta x)^2}{x} + \frac{(\delta y)^2}{y} \right). \end{aligned} \quad [7]$$

The entropy matrix is

$$E = \begin{pmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{pmatrix} = \begin{pmatrix} \frac{1}{x} & 0 \\ 0 & \frac{1}{y} \end{pmatrix}. \quad [8]$$

The linearized relaxation equations around the steady state are

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} &= \begin{pmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{pmatrix} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \\ &= H_{ss} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \end{aligned} \quad [9]$$

in which it follows from Eqs. 4 that

$$\begin{aligned} H_{xx} &= 2\kappa_+ xy - 3\kappa_- x^2 - \text{K} \\ H_{xy} &= \kappa_+ x^2 \end{aligned} \quad [10]$$

$$H_{yx} = -2\kappa_+ xy + 3\kappa_- x^2$$

$$H_{yy} = -\kappa_+ x^2$$

Using the steady-state values from Eqs. 5 in Eqs. 10 yields

$$\begin{aligned} H_{ss} &= \begin{pmatrix} \text{K} - \kappa_- \text{R}_Y^2 & \kappa_+ \text{R}_Y^2 \\ \kappa_- \text{R}_Y^2 - 2\text{K} & -\kappa_+ \text{R}_Y^2 \end{pmatrix} \\ &\equiv \begin{pmatrix} \text{K} - \beta_- & \beta_+ \\ \beta_- - 2\text{K} & -\beta_+ \end{pmatrix}. \end{aligned}$$

The second equality defines β_\pm by $\beta_\pm \equiv \kappa_\pm \text{R}_Y^2$.

The stability of this linear regime around the steady state is determined by the eigenvalues of H_{ss} . These are denoted by λ_\pm and are

$$\begin{aligned} \lambda_\pm &= -\frac{1}{2} (\beta_- + \beta_+ - \text{K}) \\ &\quad \pm \frac{1}{2} \sqrt{(\beta_- + \beta_+ - \text{K})^2 - 4\beta_+ \text{K}}. \end{aligned} \quad [12]$$

Whenever $\beta_- + \beta_+ > \text{K}$, it follows that the steady state is stable. There are two possibilities: (i) two distinct, negative, real eigenvalues if $(\beta_- + \beta_+ - \text{K})^2 > 4\beta_+ \text{K}$; and (ii) two complex conjugate roots with identical negative real parts if $(\beta_- + \beta_+ - \text{K})^2 \leq 4\beta_+ \text{K}$. This concludes the traditional linear stability analysis.

Eq. 7 makes it clear that

$$(\delta^2 S)_{ss} \leq 0. \quad [13]$$

If $(\delta^2 S)_{ss}$ is to serve as a Liapunov function, its time derivative must be ≥ 0 along the trajectories of the solutions to the rate equations. By using Eqs. 7 and 9, it follows that

$$\begin{aligned} \frac{d}{dt} (\delta^2 S)_{ss} &= -\text{K}_B \frac{d}{dt} \left[(\delta x, \delta y) E_{ss} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \right] \\ &= -\text{K}_B (\delta x, \delta y) [H_{ss}^\dagger E_{ss} + E_{ss} H_{ss}] \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}; \end{aligned} \quad [14]$$

with Eqs. 5, 8, 9, and 10, we get

$$H_{ss}^\dagger E_{ss} + E_{ss} H_{ss} = \frac{1}{\text{R}_Y} \begin{pmatrix} 2(\text{K} - \beta_-) & \frac{\beta_+(2\beta_- - \text{K})}{\beta_- + \text{K}} \\ \frac{\beta_+(2\beta_- - \text{K})}{\beta_- + \text{K}} & -2 \frac{\beta_+^2}{\beta_- + \text{K}} \end{pmatrix}. \quad [15]$$

Because

$$\det(H_{ss}^\dagger E_{ss} + E_{ss} H_{ss}) = \frac{4\beta_-^2 + \beta_- \text{K} - 5\beta_+^2 \text{K}^2}{\text{R}_Y^2 (\beta_- + \text{K})^2}, \quad [16]$$

it is ≥ 0 whenever $4\beta_- \geq 5\text{K}$, and it is ≤ 0 whenever $4\beta_- \leq 5\text{K}$. In the latter case, this leads to $d/dt (\delta^2 S)_{ss} \leq 0$ for appropriate choices of δx and δy . The requirement that the roots in Eq. 12 confer stability is $\beta_- + \beta_+ > \text{K}$, whereas the failure of $(\delta^2 S)_{ss}$ as a Liapunov function requires that $4\beta_- \leq 5\text{K}$. Both requirements may be met simultaneously by a measurable set of parameter values for β_+ , β_- , and K (Fig. 1). There is also a special subcase of this domain satisfying

$$(\beta_- + \beta_+ - \text{K})^2 = 4\beta_+ \text{K}. \quad [17]$$

With the α from Eqs. 5, for fixed K and fixed α , the choices for β_+ are reduced to the two values

$$\beta_{+(\pm)} = \frac{3 + \alpha}{(1 + \alpha)^2} \text{K} (\pm) \frac{\sqrt{8 + 4\alpha}}{(1 + \alpha)^2} \text{K}. \quad [18]$$

Eq. 12 shows that, for these choices, the relaxation eigenvalues, λ_\pm , are now "doubly degenerate" (i.e., equal) with the value

$$\lambda_\pm = -\frac{1}{2} (\beta_- + \beta_+ - \text{K}) = -\sqrt{\text{K} \beta_{+(\pm)}}. \quad [19]$$

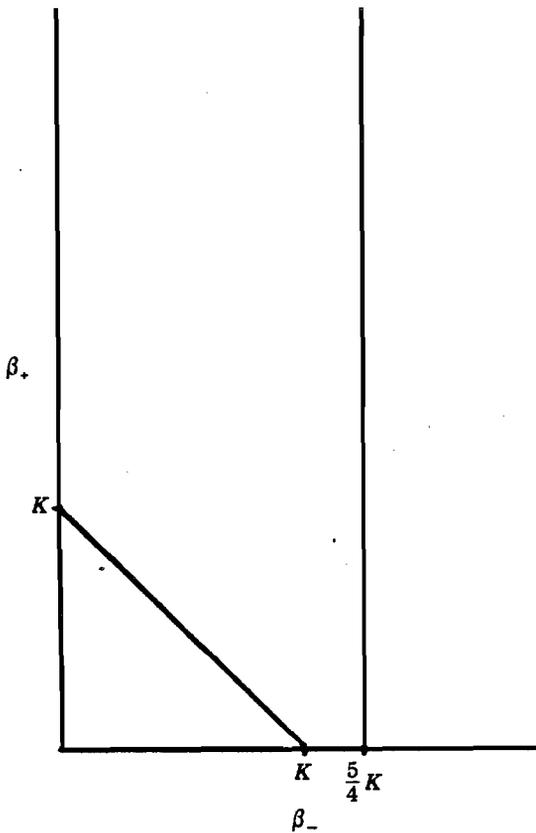


FIG. 1. $\beta_- + \beta_+ > K$ and $4\beta_- \leq 5K$.

In the linear vicinity of the stable steady states, an initial deviation from the steady-state values, at time $t = 0$ and denoted by $[\delta x(0), \delta y(0)]$, will evolve in time according to a solution of Eq. 9. There exists a similarity transformation of $H_{ss} \rightarrow S^{-1}H_{ss}S$ such that

$$S^{-1}H_{ss}S = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \quad [20]$$

with S given by

$$\begin{pmatrix} \beta_+ + \lambda_+ & \beta_+ + \lambda_- \\ \beta_- - 2K & \beta_- - 2K \end{pmatrix} \\ \det S = (\beta_- - 2K)(\lambda_+ - \lambda_-) \quad [21]$$

and

$$S^{-1} = \begin{pmatrix} \beta_- - 2K & -(\beta_+ + \lambda_-) \\ -(\beta_- - 2K) & \beta_+ + \lambda_+ \end{pmatrix} \frac{1}{\det S}$$

The solutions to Eq. 9 can always be written

$$\begin{pmatrix} \delta x(t) \\ \delta y(t) \end{pmatrix} = S \begin{pmatrix} e^{\lambda_+ t} & 0 \\ 0 & e^{\lambda_- t} \end{pmatrix} S^{-1} \begin{pmatrix} \delta x(0) \\ \delta y(0) \end{pmatrix} \quad [22]$$

This means that along the trajectories of the solutions for the rate equations

$$\frac{d}{dt} (\delta^2 S)_{ss} = -K_B [\delta x(0), \delta y(0)] (S^{-1})^\dagger \begin{pmatrix} e^{\lambda_+ t} & 0 \\ 0 & e^{\lambda_- t} \end{pmatrix} S^\dagger \\ \times (H_{ss}^\dagger E_{ss} + E_{ss} H_{ss}) S \begin{pmatrix} e^{\lambda_+ t} & 0 \\ 0 & e^{\lambda_- t} \end{pmatrix} S^{-1} \begin{pmatrix} \delta x(0) \\ \delta y(0) \end{pmatrix} \quad [23]$$

At $t = 0$, this expression reduces to the $t = 0$ value of Eq. 14. The invalidity of using $(\delta^2 S)_{ss}$ as a Liapunov function around

the steady state is seen at the initial time, $t = 0$, when Eq. 23 satisfies

$$\frac{d}{dt} (\delta^2 S)_{ss} |_{t=0} = -K_B \frac{1}{R_y} \left(2(K - \beta_-) [\delta x(0)]^2 - 2 \frac{\beta_+^2}{\beta_- + K} \right. \\ \left. \times [\delta y(0)]^2 + 2\beta_+ \frac{2\beta_- - K}{\beta_- + K} \delta x(0) \delta y(0) \right) < 0. \quad [24]$$

If the values of β_+ , β_- , and K are confined as shown in Fig. 1, then there is always a measurable family of $[\delta x(0), \delta y(0)]$ values satisfying Eq. 24. In the special case of λ_\pm roots which are complex conjugates (i.e., damped oscillations), Eq. 12 yields the frequency ν when $4\beta_+ K > (\beta_- + \beta_+ - K)^2$

$$\nu = \frac{1}{2} \sqrt{4\beta_+ K - (\beta_- + \beta_+ - K)^2} \quad [25]$$

For every integer n , the time-dependent factors in Eq. 23 satisfy

$$e^{\lambda_{\pm} t_n} = e^{-\frac{1}{2}(\beta_+ + \beta_- - K)t_n}$$

where

$$t_n \equiv n 2\pi \frac{1}{\nu} \quad [26]$$

This means that for every integer multiple of the oscillation period, $2\pi/\nu$, we have

$$\frac{d}{dt} (\delta^2 S)_{ss} |_{t=t_n} = e^{-(\beta_+ + \beta_- - K)t_n} \frac{d}{dt} (\delta^2 S)_{ss} |_{t=0} < 0 \quad [27]$$

whenever Eq. 24 is satisfied. In the special subcase given by Eq. 17, Eq. 23 reduces to

$$\frac{d}{dt} (\delta^2 S)_{ss} = -\frac{K_B}{R_y} e^{-(\beta_+ + \beta_- - K)t} \left(2(K - \beta_-) [\delta x(0)]^2 \right. \\ \left. - 2 \frac{\beta_+^2}{\beta_- + K} [\delta y(0)]^2 \right. \\ \left. + 2\beta_+ \frac{2\beta_- - K}{\beta_- + K} \delta x(0) \delta y(0) \right) < 0 \quad [28]$$

for all $(t) > 0$ wherever Eq. 24 is satisfied.

These deficiencies in $(\delta^2 S)_{ss}$ as a Liapunov function are not shared by the choice of a Liapunov function determined by the covariance of the steady-state fluctuations as advocated by Keizer (5) and Fox (1). When the latter is constructed for the example used here, a *bona fide* Liapunov function is obtained. Most importantly, it is seen to involve explicitly the input parameters K and R_y . These quantities are not in E_{ss} or in E generally, because the local entropy is determined by equilibrium behavior only. At steady states, both the fluctuation correlations and the Liapunov properties must depend upon the driving terms, the input parameters K and R_y .

CONCLUDING REMARKS

Because $(\delta^2 S)_{ss}$ does not work as a Liapunov function around steady states, the critical remarks by Nicolis and Prigogine (7) lose their force entirely. Nevertheless, it is necessary to address two additional remarks made by them in paragraphs seven and eight of their criticism of Fox (1).

Their reference to Lax and Maxo (references 9 and 10 cited in ref. 7) suggests that the Liapunov function based upon the covariance of the steady-state fluctuations, as proposed by Keizer, was already anticipated by Lax. In Lax' classic paper, there does appear a derivation of the fluctuation-dissipation relation for steady states, as is well known; but no connection

with or discussion of Liapunov properties appears anywhere in the paper. Nothing more is added by Mazo's paper.

The suggestion that Fox's approach suffers from "drawbacks" regarding the behavior of fluctuations at bifurcations or critical points is entirely incorrect. Fox (9) has shown how to treat such cases explicitly in a recent paper in which he derives Keizer's theory from a master equation basis. The critical fluctuations are indeed non-Gaussian, but the theory is more than adequate to describe them explicitly and quantitatively. In ref. 9 one will find other references regarding the treatment of critical points and bifurcations.

In their original criticism, Keizer and Fox (8) could only point to the failure of $(\delta^2 S)_{ss}$ as a Liapunov function. We did not then have an alternative proposal. Now, however, Keizer's theory of nonequilibrium thermodynamics at steady states (4-6)

provides an alternative proposal, already shown to be free from such mathematical deficiencies.

1. Fox, R. F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2114-2117.
2. Prigogine, I. (1973) in *The Physicist's Conception of Nature*, ed. Mehra, J. (Reidel, Boston), pp. 585-588.
3. Onsager, L. & Machlup, S. (1953) *Phys. Rev.* **91**, 1505-1512.
4. Keizer, J. (1976) *J. Chem. Phys.* **65**, 4431-4444.
5. Keizer, J. (1978) *J. Chem. Phys.* **69**, 2609-2620.
6. Keizer, J. (1979) *Acc. Chem. Res.* **12**, 243-249.
7. Nicolis, G. & Prigogine, I. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6060-6061.
8. Keizer, J. & Fox, R. F. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 192-196.
9. Fox, R. F. (1979) *J. Chem. Phys.* **70**, 4660-4663.

Comment on "Inconsistency between the Boltzmann distribution for relativistic free particles and the Planck spectrum for thermal radiation in quantum theory"

C. H. Braden, R. F. Fox, and H. A. Gersch

School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 28 September 1979)

Contrary to the result announced in the title of this paper, we find that Einstein's results for the consistency between the Planck and Boltzmann laws for atoms in thermal equilibrium hold as well for relativistic atoms.

In his paper¹ Boyer extends Einstein's 1917 derivation of the Planck radiation law² to relativistic atom velocities. In the first part of his paper Einstein showed Planck's law to be a consequence of the condition that the equilibrium distribution of atoms of large mass and negligible velocities over their internal energy states (given by the Boltzmann distribution) must be established solely by absorption and emission of radiation. This classic, beautifully simple derivation is well known, appearing in textbooks in quantum mechanics.³

Boyer's calculation relates to the second part of Einstein's paper, which investigates the motion of the atoms under the influence of radiation. In this much more delicate and intricate problem it is required that the mean square speed which atoms acquire because of their interaction with radiation of temperature T be equal to that predicted by the equipartition law of statistical mechanics. Einstein demonstrates that this requirement is fulfilled, for heavy atoms moving slowly, only if the radiation density is given by Planck's law. It is this latter calculation (as refined by Milne⁴) which Boyer extends to relativistic atom velocities and which leads to the inconsistency announced in the title of his paper.

If true, such an inconsistency would have far-reaching consequences, requiring, for example, fundamental modification in theories of early evolution of the universe.⁵ One should also see the inconsistency between the Planck and Boltzmann laws for relativistic atoms by extending Einstein's simple, classic derivation of Planck's law given in the first part of his paper.

In this comment we review the simple derivation of the Planck radiation law using quantum

field theory,⁶ which amounts to giving a modern version of Einstein's work. Inspection of the details of this derivation shows it is in no way restricted to nonrelativistic atom motions, but rather indicates consistency between the Planck and Boltzmann laws for any atom velocity.

We suppose the system to contain atoms which may be in states A or B and a radiation field of photons γ which can freely exchange energy by the reversible process

$$A \rightleftharpoons B + \gamma. \quad (1)$$

Atom states A and B will in general describe center-of-mass motion as well as internal electron motion. Thus, for example, a typical state A for a center-of-mass momentum \vec{p} and electronic state ψ is $|A\rangle = |e^{i\vec{p}\cdot\vec{R}}\psi\rangle$, where \vec{R} is the center-of-mass coordinate.

Let $N(A)$ and $N(B)$ be the number of atoms in states A and B , respectively, when equilibrium is established at some temperature T . If P_{γ} and P_a are, respectively, the transition probabilities per sec for the forward ($A \rightarrow B + \gamma$) and reverse ($B + \gamma \rightarrow A$) reactions, the condition for equilibrium is

$$N(B)P_a = N(A)P_{\gamma}. \quad (2)$$

The equilibrium ratio of numbers of atoms in states B and A is given by Boltzmann's law:

$$\frac{N(B)}{N(A)} = \frac{e^{-E_B/\hbar T}}{e^{-E_A/\hbar T}}, \quad (3)$$

where E_A and E_B are the energies of atoms in states A and B , and include both center-of-mass energy and internal energy. Then Planck's radiation law, expressed in quantum language, is that $n_{\mathbf{k},\alpha}$, the number of photons of wave vector

\vec{k} ($|k| = \omega/c$) and polarization state α in thermal equilibrium, depends only on frequency ω in the form

$$n_{k,\alpha} = (e^{\hbar\omega/kT} - 1)^{-1}, \quad (4)$$

with frequency ω related to the difference in energy between states A and B :

$$\hbar\omega = E_A - E_B. \quad (5)$$

Transition probabilities are calculated in first-order, time-dependent perturbation theory with atom-photon interaction H' ,

$$H' = \frac{-e}{m_e c} \sum_i \vec{p}_i \cdot \vec{A}, \quad (6)$$

where m_e is the electron rest mass, \vec{p}_i is the momentum of the i th electron, and \vec{A} is the vector potential for the radiation field. The obvious objection to the form of H' being a nonrelativistic expression for the electron-photon interaction is really not relevant, since all we require is the ratio P_e/P_a . This ratio, given by Eq. (13) below, is independent of the matrix elements of H' between states A and B , and the same result for this ratio is obtained by using a relativistic form for H' .

The vector potential \vec{A} expressed in terms of annihilation ($a_{k,\alpha}$) and creation ($a_{k,\alpha}^\dagger$) operators for photons of wave vector \vec{k} and polarization state α is

$$\vec{A}(\vec{r}, t) = \left(\frac{2\pi c^2 \hbar}{\Omega} \right)^{1/2} \sum_{k,\alpha} \hat{\epsilon}_{k,\alpha} \omega_k^{-1/2} \times (a_{k,\alpha} e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} + a_{k,\alpha}^\dagger e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)}), \quad (7)$$

where $\hat{\epsilon}_{k,\alpha}$ is a unit vector in the polarization direction. Straightforward application of first-order, time-dependent perturbation theory then gives

$$P_e = \frac{e^2}{\hbar \tilde{m}^2 c^3} (n_{k,\alpha} + 1) \omega \sum_i |\langle B | e^{-i\vec{k} \cdot \vec{r}_i} \hat{\epsilon}_{k,\alpha} \cdot \vec{p}_i | A \rangle|^2, \quad (8)$$

$$P_a = \frac{e^2}{\hbar \tilde{m}^2 c^3} n_{k,\alpha} \omega \sum_i |\langle A | e^{i\vec{k} \cdot \vec{r}_i} \hat{\epsilon}_{k,\alpha} \cdot \vec{p}_i | B \rangle|^2, \quad (9)$$

where photon frequency ω agrees with Eq. (5),

$$\hbar\omega = E_A - E_B. \quad (10)$$

The photon momentum $\hbar \vec{k}$ satisfies the momentum conservation,

$$\vec{p} = \vec{p}' + \hbar \vec{k}, \quad (11)$$

with \vec{p} (\vec{p}') the center-of-mass momentum of state A (B), as follows from integrating matrix

elements in Eqs. (8) and (9) over center-of-mass coordinate R , setting the i th electron coordinate $\vec{r}_i = \vec{R} + \vec{\rho}_i$, with $\vec{\rho}_i$ the relative electron coordinate.

Because of the Hermitian character of the perturbation matrices which determine these transition probabilities,

$$\begin{aligned} \langle B | e^{-i\vec{k} \cdot \vec{r}_i} \hat{\epsilon}_{k,\alpha} \cdot \vec{p}_i | A \rangle &= \langle A | \vec{p}_i \cdot \hat{\epsilon}_{k,\alpha} e^{i\vec{k} \cdot \vec{r}_i} | B \rangle \\ &= \langle A | e^{i\vec{k} \cdot \vec{r}_i} \hat{\epsilon}_{k,\alpha} \cdot \vec{p}_i | B \rangle^*, \end{aligned} \quad (12)$$

the ratio of the transition probabilities is independent of these matrix elements, depending only on photon occupation numbers:

$$\frac{P_e}{P_a} = \frac{n_{k,\alpha} + 1}{n_{k,\alpha}}. \quad (13)$$

The condition for equilibrium expressed by Eq. (2) implies that the ratio of transition probabilities equals the ratio of numbers of atoms in the two states A and B :

$$\frac{P_e}{P_a} = \frac{N(B)}{N(A)} = \frac{e^{-E_B/kT}}{e^{-E_A/kT}}. \quad (14)$$

Combining Eqs. (13) and (14) we obtain Planck's radiation law:

$$n_{k,\alpha} = (e^{\hbar\omega/kT} - 1)^{-1}. \quad (15)$$

Compared with Einstein's 1917 derivation, this standard quantum-mechanical derivation differs in three respects: (a) spontaneous and induced emissions appear automatically from the boson character of photons, (b) detailed balance between forward and reverse reactions is an automatic consequence of the Hermitian interaction between atoms and radiation field, and (c) no restriction on atom velocities (described by the states $|A\rangle$ and $|B\rangle$) has been made. For relativistic velocities, in order to treat the atom as a "particle" having certain discrete energies allowed by quantum mechanics, one considers those energy levels to be levels of proper energy, since the only privileged frame of reference is the atom itself. Thus we can assign a proper mass M to the level A , and a smaller mass M' to the lower level B . This only changes the detailed form of E_A and E_B , $E_A = (M^2 c^4 + c^2 p^2)^{1/2}$, $E_B = (M'^2 c^4 + c^2 p'^2)^{1/2}$, but does not affect the conservation laws $\hbar\omega = E_A - E_B$ and $\vec{p} = \vec{p}' + \hbar \vec{k}$. Of course the photon frequency depends on M, M', \vec{p} , and the direction of \vec{k} relative to \vec{p} . From the conservation laws, one easily finds

$$\hbar\omega = \frac{c^4 (M^2 - M'^2)}{2 [(M^2 c^4 + c^2 p^2)^{1/2} - c \vec{p} \cdot \hat{k}]},$$

where \hat{k} is a unit vector in the direction of photon

propagation. This can be written, using $M\gamma c^2 = (M^2 c^4 + c^2 p^2)^{1/2}$, $(M - M')c^2 = E_0 = \hbar \omega_0$ and $\vec{p} = M\gamma \vec{v}$, as

$$\hbar \omega = \hbar \omega_0 \frac{1}{(1 - \vec{v} \cdot \vec{k}/c)} \frac{(M + M')}{2M}.$$

Reference 1 mistakenly assigns photons in the rest frame of the atoms (assumed the same for atoms in both upper and lower states) to have the frequency ω_0 , which amounts to omission of the factor $(M + M')/2M$ and thereby introduces an error in the analysis except in the limit $M = M'$ or $\omega_0 = 0$.

Thus we conclude that Einstein's original simple derivation of the consistency between Planck and Boltzmann laws for atoms in thermal equilibrium with radiation holds as well for relativistic atoms. Note, however, that the Planck and Boltzmann laws are not form invariant under a Lorentz transformation. The usual form for the laws is exhibited only in the zero-momentum frame of the system of atoms and radiation which are in thermal equilibrium.

We thank Dr. Henry Valk for calling Boyer's paper to our attention.

¹T. H. Boyer, *Phys. Rev. D* **19**, 3635 (1979).

²A. Einstein, *Phys. Z.* **18**, 121 (1917).

³See, for example, L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968) and E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970).

⁴E. A. Milne, *Proc. Cambridge Philos. Soc.* **23**, 465 (1926).

⁵S. Weinberg, *The First Three Minutes* (Basic Books,

New York, 1977).

⁶Our derivation follows that given by J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, Massachusetts, 1970). The first demonstration of consistency between Planck and Boltzmann laws for relativistic atoms was given by P. A. M. Dirac, *Proc. R. Soc. London* **A114**, 243 (1927).

7

The Origin of Irreversibility in Quantum Statistical Mechanics

Ronald Forrest Fox, School of Physics

Georgia Institute of Technology

October, 1980

To appear in *Advances in Mathematics Supplement*

Introduction

During the fall of 1964, I began the first year of my graduate studies at the California Institute of Technology. Shortly after the term began, I came across the book:¹ Probability and Related Topics in Physical Sciences, by Mark Kac. This book also contained three sets of special lectures by Uhlenbeck, Hibbs, and van der Pol. About the contributions by Uhlenbeck, Kac wrote in the preface:

"The third (and longest!) is Probability in Some Problems of Classical Statistical Mechanics. This chapter is supplemented by two lectures (reproduced in Appendix I) by Professor G. E. Uhlenbeck on the Boltzmann equation. It was indeed fortunate that we were able to include in this volume so lucid an exposition of the physical ideas underlying this difficult and important field and to secure the collaboration of the foremost authority on these questions!

Our Chapter III is to a large extent a running commentary on some of the points raised in Uhlenbeck's first lecture and although the chapter is self-contained, the reader is strongly advised to read one in conjunction with the other. Uhlenbeck's second lecture brings the reader to the frontiers of our knowledge and opens a vast field where opportunities for further contributions are almost unlimited.

Our debt to Uhlenbeck is, however, infinitely greater. Almost all of Chapter III has been directly or indirectly influenced by discussions and correspondence stretching over almost fifteen years of friendship and scientific association."

I soon learned that Kac and Uhlenbeck had just the year before moved to the Rockefeller University where Detlev Bronk was building a new mathematical

and physical sciences addition to the University's traditional, outstanding medical research enterprise. By the end of the 1960's Bronk built programs in mathematical physics and in philosophy which were internationally recognized as preeminent in these specialties. Kac and Uhlenbeck directed the efforts in mathematics and statistical physics.

Late in the fall of 1964, I applied to Bronk for transfer to Rockefeller University in order to study with Kac and Uhlenbeck. I knew that the student body at Rockefeller was very small, and imagined that I might be able to work closely with these two men. The culmination of the usual sequence of procedures for admission was a personal interview with Bronk. As soon as I entered his office, Bronk extolled my great good sense in having recognized the attractiveness of working under Kac and Uhlenbeck, and granted me admission to the graduate school on the spot! It is impossible to overestimate the benefit to me to have been allowed to study from 1965 to 1969 with Kac and Uhlenbeck. My thesis, "Contributions to the Theory of Non-Equilibrium Thermodynamics", resulted from a very active collaboration with both men.

The present paper reports subsequent work in quantum statistical mechanics. Even though all of this work was done after I left Rockefeller, the influence of Kac and Uhlenbeck is still very strong. The basic underlying philosophy has been beautifully expounded by Uhlenbeck in Chapter 1 of Lectures in Statistical Mechanics². The methodology chosen to implement this philosophy, however, follows from a technique which Kac has so successfully used in both mathematics and physics. The technique is to study a model system with two properties: 1) enough simplicity to be tractable; and 2) enough richness to illuminate the "real" problem. This paper uses a model to explain the origin of irreversibility in quantum statistical mechanics.

Contraction of the Description

In classical physics, the relationships between the time reversible microscopic dynamics and the time-irreversible macroscopic dynamics has been understood since the days of Boltzmann and Gibbs.² The microscopic picture for N classical, interacting particles is given by a $6N$ -dimensional phase space, called Γ -space, in which the microscopic state of the N particle system is represented as a single $6N$ -dimensional point. The time evolution of the system is represented by a trajectory of points in Γ -space. The trajectory is the solution to Newton's equations of motion. For this reason, it is known that this description is time reversible. Boltzmann observed that for N particle, interacting systems the Γ -trajectory is chaotic, although strictly continuous and differentiable. Boltzmann associated measureable regions of Γ -space with macroscopic states of the system. One region was overwhelmingly the largest and Boltzmann identified it as the equilibrium region. This means that for almost all Γ -points in the equilibrium region, the macroscopic variables have their equilibrium values. The non-equilibrium regions of Γ -space are individually much smaller than the equilibrium region, and they are of great variety. A Γ -point initially located in a non-equilibrium region of Γ -space will surely go into the equilibrium region, where it will almost surely remain because this region is so large. The macroscopic variables reflect this asymmetric time course by changing from non-equilibrium values into the equilibrium values. This is Boltzmann's argument for irreversibility in the macroscopic dynamics.²

Boltzmann's ideas were vigorously attracted by Zermelo³, who invoked Poincare's recurrence theorem, "Almost every Γ -point will reappear in an arbitrary small sphere centered at its initial position at t_0 , the time origin."

Thus, a Γ -point initially in a non-equilibrium region will certainly come back to that region even though a lot of intervening time is spent in the equilibrium region. This fact undoes the apparent irreversibility. Boltzmann replied that the recurrence of the Γ -point merely gave rise to fluctuations in the macroscopic variables, and that most of the time the Γ -point would be found in the equilibrium region; but in vain.

Gibbs had a view of Γ -space which made Boltzmann's ideas manifest. He studied the behavior of an "ensemble" of Γ -points. An ensemble of Γ -points moves like an incompressible fluid in $6N$ -dimensions. Even though its shape will radically change in time, its hypervolume is invariant. For an ensemble, it is true that almost every Γ -point exhibits recurrence, but they do so at immensely many different recurrence times. Therefore, at any single time, the vast majority of the ensemble's Γ -points are in the equilibrium region of Γ -space. Associating the ensemble with the macroscopic description yields macroscopic variables which fluctuate and exhibit irreversible behavior.

Appropriately chosen ensembles generate macroscopic descriptions which utilize relatively few macroscopic variables as compared with the $6N$ microscopic variables. This reduction in the number of variables is called the "contraction of the description".²

Reduced Density Matrices

It is desirable to translate the Boltzmann-Gibbs-Uhlenbeck explanation of the origin of irreversibility in classical physics into its analogue in quantum physics. On the one hand, we are immediately confronted with the difficulty of translating the idea of Γ -space into quantum mechanics. On the other hand we must avoid the ambiguities of "coarse-graining" Γ -space, which the classical ideas imply. It turns out that both of these dilemmas are

avoided by taking a slightly altered perspective which is natural in the quantum mechanical setting. The solution is to employ a density matrix formalism and to achieve a "contraction of the description" by generating "reduced" density matrices.

The density matrix⁴ provides an alternative to the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \psi = H\psi \quad (1)$$

in the form

$$i\hbar \frac{\partial}{\partial t} \rho = [H, \rho] \quad (2)$$

H is the Hamiltonian. The right-hand side of (2) denotes the commutator of H and ρ

$$[H, \rho] = H\rho - \rho H \quad (3)$$

in which ρ is the density "matrix", or "operator", which is related to ψ by

$$\rho \equiv |\psi\rangle\langle\psi|$$

In (4), $\langle\psi|$ denotes the Schrödinger wave function, ψ , in Dirac's ket notation.

One way to implement the idea of "contraction of the description" while dealing with density matrices is to construct "reduced" density matrices. In order to illustrate this technique, suppose that we divide up a system into two parts: a subsystem, S , and the remainder of the system, R . We may associate a Hamiltonian H_S with the subsystem by itself, and a Hamiltonian H_R with the remainder by itself. Denote the eigenkets of H_S and H_R respectively by $|\overset{S}{r}\rangle$ and $|\overset{R}{s}\rangle$. The product kets, $|r\rangle|s\rangle$, span the Hilbert space of the complete system, and the original Hamiltonian for the complete system may be written

$$H = H_S \otimes 1_R + 1_S \otimes H_R + H_I \tag{5}$$

in which H_I includes all interactions between the subsystem S and the remainder R. If we wish to follow the time development of S while ignoring the details of R, we may use the reduced density matrix, ρ_{Red} , defined by

$$\begin{aligned} \rho_{Red} &\equiv \sum_{|r\rangle} \langle r | \rho | r \rangle \\ &\equiv \text{Trace}_R \rho \end{aligned} \tag{6}$$

Operator Calculus

In this section, several operator calculus results are summarized for subsequent use.

To solve (2), it is convenient to introduce the "commutator operator", $[H, \cdot]$, defined by

$$[H, \cdot]M = HM - MH \tag{7}$$

in which M is an arbitrary operator. While M acts in the original Hilbert space, $[H, \cdot]$ acts in the space of operators. It is called a "superoperator".

The solution to (2) can be written

$$\rho(t) = \exp \left\{ -\frac{i}{\hbar} t [H, \cdot] \right\} \rho(o) \tag{8}$$

in which $\rho(o)$ is the initial value density matrix (operator), and the exponentiated superoperator is itself a superoperator which acts on $\rho(o)$. An easily proved identity, for arbitrary M, is

$$\exp \left\{ -\frac{i}{\hbar} t [H, \cdot] \right\} M = \exp \left\{ -\frac{i}{\hbar} t H \right\} M \exp \left\{ \frac{i}{\hbar} t H \right\} \tag{9}$$

The utility of the reduced density matrix is enhanced when the "interaction

picture" is used. Instead of solving (2) directly, as in (8), we look at σ , defined by

$$\rho(t) \equiv \exp \left\{ -\frac{i}{\hbar} t [H_S \otimes 1_R + 1_S \otimes H_R, \cdot] \right\} \sigma(t) \quad (10)$$

in which (5) has been used. This yields an equation⁵ for $\sigma(t)$:

$$i\hbar \frac{\partial}{\partial t} \sigma(t) = [\hat{H}_I(t), \cdot] \sigma(t) \quad (11)$$

in which $\hat{H}_I(t)$ is defined by

$$\hat{H}_I(t) = \exp \left\{ \frac{i}{\hbar} t (H_S \otimes 1_R + 1_S \otimes H_R) \right\} H_I \exp \left\{ -\frac{i}{\hbar} t (H_S \otimes 1_R + 1_S \otimes H_R) \right\} \quad (12)$$

The solution to (11) is not as easy to express as was the solution, (8), to (2); because of the t -dependence in $\hat{H}_I(t)$. For two different times, t and s , we find

$$[\hat{H}_I(t), \hat{H}_I(s)] \neq 0 \quad (13)$$

because $[H_S \otimes 1_R, H_I] \neq 0$ and $[1_S \otimes H_R, H_I] \neq 0$.

We may overcome the commutativity problem expressed in (13) by introducing the "t-ordered exponential". Consider the general vector equation

$$\frac{d}{dt} \vec{a} = M(t) \vec{a} \quad (14)$$

in which \vec{a} is an N -component vector and $M(t)$ is a t -dependent operator ($N \times N$ matrix) acting in the vector space in which \vec{a} is found. The solution to (14), with initial value vector \vec{a}_0 , is

$$\begin{aligned} \vec{a}(t) = & \vec{a}_0 + \int_0^t ds M(s) \vec{a}_0 + \int_0^t ds_1 \int_0^{s_1} ds_2 M(s_1) M(s_2) \vec{a}_0 + \dots \\ & + \int_0^t ds_1 \int_0^{s_1} ds_2 \dots \int_0^{s_{n-1}} ds_n M(s_1) M(s_2) \dots M(s_n) \vec{a}_0 + \dots \end{aligned} \quad (15)$$

$$\begin{aligned}
 &= (1 + \sum_{n=1}^{\infty} \int_0^t ds_1 \int_0^{s_1} ds_2 \dots \int_0^{s_{n-1}} ds_n M(s_1) M(s_2) \dots M(s_n)) \vec{a}_0 \\
 &\equiv \overleftarrow{T} \exp \left\{ \int_0^t ds M(s) \right\} \vec{a}_0
 \end{aligned}$$

The third equality defines the symbol $\overleftarrow{T} \exp \left\{ \int_0^t ds M(s) \right\}$, which must not be confused with the ordinary exponential, $\exp \left\{ \int_0^t ds M(s) \right\}$. A reverse t-ordered exponential also exists:

$$\overleftarrow{T} \exp \left\{ \int ds M(s) \right\} \equiv 1 + \sum_{n=1}^{\infty} \int_0^t ds_1 \int_0^{s_1} ds_2 \dots \int_0^{s_{n-1}} ds_n M(s_n) \dots M(s_2) M(s_1) \tag{16}$$

It may be used to solve the vector equation

$$\frac{d}{dt} \vec{a}^\dagger = \vec{a}^\dagger M(t) \tag{17}$$

in which \dagger denotes "adjoint". It also provides the inverse of a forward t-ordered exponential:

$$\overleftarrow{T} \exp \left\{ - \int_0^t ds M(s) \right\} = \left(\overleftarrow{T} \exp \left\{ \int_0^t ds M(s) \right\} \right)^{-1} \tag{18}$$

We may now write the solution to (11) in the form

$$\sigma(t) = \overleftarrow{T} \exp \left\{ - \frac{i}{\hbar} \int_0^t ds [\hat{H}_I(s), \cdot] \right\} \sigma(o) \tag{19}$$

For arbitrary $M(s)$, the analogue of (9) may be proved⁵:

$$\overleftarrow{T} \exp \left\{ - \frac{i}{\hbar} \int ds [\hat{H}_I(s), \cdot] \right\} M = \overleftarrow{T} \exp \left\{ - \frac{i}{\hbar} \int ds \hat{H}_I(s) \right\} M \overleftarrow{T} \exp \left\{ \frac{i}{\hbar} \int ds \hat{H}_I(s) \right\} \tag{20}$$

Because $H_S \otimes 1_R$ and $1_S \otimes H_R$ act in orthogonal subspaces of the complete Hilbert space, (10) can be written

$$\rho(t) = \exp \left\{ -\frac{i}{\hbar} t (H_S \otimes 1_R) \right\} \exp \left\{ -\frac{i}{\hbar} t (1_S \otimes H_R) \right\} \sigma(t) \times \tag{21}$$

$$\exp \left\{ \frac{i}{\hbar} t (1_S \otimes H_R) \right\} \exp \left\{ \frac{i}{\hbar} t (H_S \otimes 1_R) \right\}$$

This implies that the reduced density matrix defined in (6) can be written

$$\rho_{\text{Red.}}(t) = \exp \left\{ -\frac{i}{\hbar} t H_S \right\} \sum_{|r\rangle} \langle r | \sigma(t) | r \rangle \exp \left\{ \frac{i}{\hbar} t H_S \right\} \tag{22}$$

Notice that all operators are now acting in the subsystem Hilbert space only.

Cumulants and Stochasticity

Equation (22) does not provide a closed description for $\rho_{\text{Red.}}(t)$ in terms of itself only, but depends upon the full density matrix, $\sigma(t)$. The idea of "contraction of the description" is to get a closed description in terms of fewer variables. This means that we must find a closed description for $\rho_{\text{Red.}}(t)$. The technique for achieving this goal is exhibited below.

In the model calculations described later, we will consider R to be a reservoir with which the subsystem S interacts. At $t = 0$, the reservoir R is assumed to be in equilibrium whereas the subsystem may be prepared so that it is in any state whatever. At $t = 0$ the interaction between S and R begins and as time proceeds, S and R both evolve in a complicated way. From (21) it follows that $\rho(0) = \sigma(0)$. We assume that the initial density matrix factors and may be expressed

$$\sigma(0) = \sigma_S(0) \otimes \sigma_R^{\text{eq.}} \tag{23}$$

in which the superscript, eq., denotes equilibrium. Using (23) in (19), and then putting (19) into (22) yields

$$\rho_{\text{Red.}}(t) = \exp \left\{ -\frac{i}{\hbar} t H_S \right\} \sum_{|r\rangle} \langle r | T \exp \left\{ -\frac{i}{\hbar} \int_0^t ds [\hat{H}_I(s), \cdot] \right\} \sigma_S(0) \times \sigma_R^{\text{eq.}} | r \rangle \exp \left\{ \frac{i}{\hbar} t H_S \right\} \tag{24}$$

Define the evolution superoperator, $E(t)$, which acts on operators in the subsystem Hilbert space only, by

$$E(t) \equiv \sum_{|r\rangle} \langle r|_I \exp \left\{ -\frac{i}{\hbar} \int_0^t ds [\hat{H}_I(s), \cdot] \right\} \sigma_R^{eq} |r\rangle \quad (25)$$

Therefore, we can write (24) as

$$\rho_{\text{Red.}}(t) = \exp \left\{ -\frac{i}{\hbar} t H_S \right\} (E(t) \sigma_S(o)) \exp \left\{ \frac{i}{\hbar} t H_S \right\} \quad (26)$$

From this form of $\rho_{\text{Red.}}(t)$, it follows that

$$i \hbar \frac{\partial}{\partial t} \rho_{\text{Red.}}(t) = [H_S, \rho_{\text{Red.}}(t)] + \quad (27)$$

$$i \hbar \exp \left\{ -\frac{i}{\hbar} t H_S \right\} \left(\left(\frac{\partial}{\partial t} E(t) \right) E^{-1}(t) E(t) \sigma_S(o) \right) \exp \left\{ \frac{i}{\hbar} t H_S \right\}$$

It will be shown below that the superoperator combination $\left(\frac{\partial}{\partial t} E(t) \right) E^{-1}(t)$ can be written

$$\left(\frac{\partial}{\partial t} E(t) \right) E^{-1}(t) \equiv G(t) \quad (28)$$

Using (9) twice, and (26) once with (28), permits us to rewrite (27) as

$$i \hbar \frac{\partial}{\partial t} \rho_{\text{Red.}}(t) = [H_S, \cdot] \rho_{\text{Red.}}(t) + i \hbar \exp \left\{ -\frac{i}{\hbar} t [H_S, \cdot] \right\} G(t) \exp \left\{ \frac{i}{\hbar} t [H_S, \cdot] \right\} \rho_{\text{Red.}}(t) \quad (29)$$

This is now a manifestly closed equation for the reduced density matrix (operator) $\rho_{\text{Red.}}(t)$. The first term on the right-hand side of (29) is the only term present if the subsystem S is decoupled from R. The second term involves a succession of three superoperators and is explicitly t-dependent. This t-dependence is the price paid in order to achieve the contraction of the description. Nevertheless, it will be seen in the model calculations below that even this t-dependence can be handled appropriately.

The factorization (23) at $t = 0$ does not persist for $t > 0$ unless $H_I = 0$ in (5). When $H_I = 0$, $G(t)$ vanishes. The resulting equation contains only the first term on the right-hand side of (29), and is a special case of (2). Consequently, it is equivalent with the Schrödinger description in (1). When $H_I \neq 0$, $G(t)$ does not vanish, and it is not expressible in the form of a simple commutator operator (superoperator), i.e.,

$$\exp \left\{ -\frac{i}{\hbar} t [H_S, \cdot] \right\} G(t) \exp \left\{ \frac{i}{\hbar} t [H_S, \cdot] \right\} \neq [H_{\text{effective}}(t), \cdot] \quad (30)$$

Therefore, (29) is not a special case of (2), nor is it equivalent to (1), with an effective Hamiltonian.⁵ An effective, Hermitean Hamiltonian would imply time reversibility. We shall see that $G(t)$ is structured so that there is time irreversibility. Nevertheless, $G(t)$ will still imply conservation of total probability, just as does an effective, Hermitean Hamiltonian.

Equation (25) exhibits the reduction of an exponentiated superoperator in the total Hilbert space to an exponential superoperator in the subsystem Hilbert space. This reduction may be expressed usefully by introducing operator cumulants.^{6,7} We may do so, because it is possible to interpret the act of reduction as a kind of "averaging". For an arbitrary superoperator in the total Hilbert space, M , the "averaged" value is a superoperator in the subsystem Hilbert space with "averaging" defined by

$$\langle M \rangle \equiv \sum_{|r\rangle} \langle r | M \sigma_R^{\text{eq}} \cdot | r \rangle \quad (31)$$

The operator cumulant expansion is expressed

$$\begin{aligned} & \langle \mathbb{T} \exp \left\{ -\frac{i}{\hbar} \int_0^t ds [\hat{H}_I(s), \cdot] \right\} \rangle \equiv \quad (32) \\ & \sum_{|r\rangle} \langle r | \mathbb{T} \exp \left\{ -\frac{i}{\hbar} \int_0^t ds [\hat{H}_I(s), \cdot] \right\} \sigma_R^{eq} | r \rangle \equiv \\ & \mathbb{T} \exp \left\{ \sum_{n=1}^{\infty} \int_0^t ds G^{(n)}(s) \right\} \end{aligned}$$

Explicit expressions exist⁷ for the n^{th} operator cumulant, $G^{(n)}$, in terms of the moments of \hat{H}_I . These moments are computed using the averaging defined in (31). Comparison with (28) verifies

$$G(t) = \sum_{n=1}^{\infty} G^{(n)}(t) \quad (33)$$

The first two cumulants⁷ are

$$G^{(1)}(t) = \left\langle -\frac{i}{\hbar} [\hat{H}_I(t), \cdot] \right\rangle \quad (34)$$

and

$$G^{(2)}(t) = -\frac{1}{\hbar^2} \int_0^t ds \langle [\hat{H}_I(t), \cdot] [\hat{H}_I(s), \cdot] \rangle - \int_0^t ds G^{(1)}(t) G^{(1)}(s) \quad (35)$$

Using (31) in (34) yields

$$\begin{aligned} G^{(1)}(t) &= -\frac{i}{\hbar} \sum_{|r\rangle} \langle r | [\hat{H}_I(t), \cdot] \sigma_R^{eq} | r \rangle \quad (36) \\ &= -\frac{i}{\hbar} \sum_{|r\rangle} (\langle r | \hat{H}_I(t) \sigma_R^{eq} | r \rangle \cdot - \cdot \langle r | \sigma_R^{eq} \hat{H}_I(t) | r \rangle) \end{aligned}$$

whereas in (35) it yields

$$\begin{aligned} G^{(2)}(t) &= -\frac{1}{\hbar^2} \int_0^t ds \sum_{|r\rangle} (\langle r | \hat{H}_I(t) \hat{H}_I(s) \sigma_R^{eq} | r \rangle \cdot + \cdot \langle r | \sigma_R^{eq} \hat{H}_I(s) \hat{H}_I(t) | r \rangle \\ &\quad - \langle r | \hat{H}_I(t) \sigma_R^{eq} \cdot \hat{H}_I(s) | r \rangle - \langle r | \hat{H}_I(s) \sigma_R^{eq} \cdot \hat{H}_I(t) | r \rangle) \quad (37) \\ &\quad - \int_0^t ds G^{(1)}(t) G^{(1)}(s) \end{aligned}$$

in each of which the dots show where to place the subsystem operators upon which these superoperators act.

Notice that $G^{(1)}(t)$ acts on the subsystem Hilbert space operators just like an effective Hamiltonian, e.g. like the right-hand side of (30). $G^{(2)}(t)$, however cannot be rewritten in the same way. The higher order cumulants also behave like $G^{(2)}$ in this regard.

The relationship between the averaging defined by (31) and truly stochastic averaging may be examined by studying the "correlation functions", such as in (37):

$$C(t,s) \equiv \sum_{|r\rangle} \langle r | \hat{H}_I(t) \hat{H}_I(s) \sigma_R^{eq.} | r \rangle \quad (38)$$

$$- \sum_{|r\rangle} \langle r | \hat{H}_I(t) \sigma_R^{eq.} | r \rangle \sum_{|r'\rangle} \langle r' | \hat{H}_I(s) \sigma_R^{eq.} | r' \rangle$$

The subtracted term in (38) corresponds with part of the last term on the right-hand side of (37), and can usually be omitted because $G^{(1)} = 0$ in many model calculations. It's included here for generality. The question is whether $C(t,s)$ vanishes as $|t-s| \rightarrow \infty$? If it does, does the "correlation time", τ_c , exist?

$$\tau_c \equiv \frac{\left| \int_0^\infty dT T C(T) \right|}{\left| \int_0^\infty dT C(T) \right|} \quad (39)$$

in which $T \equiv |t-s|$. For truly stochastic processes, both of these conditions are met. For reduced density matrices, these conditions must be demonstrated explicitly for each specific model. In anticipation of the outcome of these calculations for the models discussed later in this paper, it may be claimed that the asymptotic condition is fulfilled by an application of the Riemann-Lebesgue lemma, and the correlation time exists as well. Therefore, the

reduced density matrix behaves as though it were the average of a truly stochastic process.

Characteristic Functionals

Following the stochastic analogy to its logical conclusion suggest the construction⁵ of the characteristic functional for the process. The characteristic functional for a stochastic process is the generator of moments of all orders. Functional differentiation n times generates the nth moment. Using the cumulant expressions in terms of moments⁵ provides cumulants to order n. Usually, the cumulant expansion (33) does not sum in closed form. Nevertheless, the characteristic functional can be found in closed form, in some cases. The Gaussian case has these properties,⁵ as is discussed later in this paper.

In order to construct the characteristic functional, we must introduce an "auxillary" superoperator⁵ with respect to which functional derivatives will be taken. This superoperator is denoted by K(t), and it acts in the subsystem Hilbert space.

The fundamental evolution superoperator is given in (11), and is $[\hat{H}_I(t), \cdot]$. It acts in the complete Hilbert space. To construct the characteristic functional for it, the auxillary superoperator and it must be combined. Since these two superoperators act in Hilbert spaces of different size, their combination must be approached with care. The required care is achieved by explicitly using an eigenket representation.

The eigenkets of H_S were denoted by $|S\rangle$ earlier. For the present purposes, we will also denote them by $|\mu\rangle$, $|\nu\rangle$, $|\mu'\rangle$, and $|\nu'\rangle$. Let M be an operator in the complete Hilbert space. Its matrix elements with respect to the subsystem Hilbert space are

$$M_{\mu', \nu'} \equiv \langle \mu' | M | \nu' \rangle \quad (40)$$

This matrix is still an operator in the remainder of the complete Hilbert space. We may now write the action of $[\hat{H}_I(t), \cdot]$ on M in the form

$$\begin{aligned} [\hat{H}_I(t), \cdot] M &= \hat{H}_I(t) M - M \hat{H}_I(t) \quad (41) \\ &= |\mu\rangle\langle\mu| \hat{H}_I(t) |\mu'\rangle\langle\mu'| M |\nu\rangle\langle\nu| - |\mu\rangle\langle\mu| M |\nu'\rangle\langle\nu'| \hat{H}_I(t) |\nu\rangle\langle\nu| \end{aligned}$$

in which summation over repeated labels is implicit. This implies that the superoperator, $[\hat{H}_I(t), \cdot]$, has subsystem "matrix" (tetratic) elements given by

$$([\hat{H}_I(t), \cdot])_{\mu\nu\mu'\nu'} \equiv \quad (42)$$

$$\delta_{\nu\nu'} \langle\mu| \hat{H}_I(t) |\mu'\rangle \dots \delta_{\mu\mu'} \langle\nu'| \hat{H}_I(t) |\nu\rangle$$

The dots indicate the location of operators in the remainder Hilbert space when this superoperator acts in the remainder Hilbert space, in parallel with the explicit case in (41). The auxiliary superoperator, $K(t)$, can also be given explicit labels for the subsystem Hilbert space: $K(t) \rightarrow K_{\mu\nu\mu'\nu'}(t)$.

We now form the subsystem scalar combination of $K(t)$ and $[\hat{H}_I(t), \cdot]$, which is still a superoperator in the remainder Hilbert space:

$$L(t) \equiv \sum_{\substack{|\mu\rangle|\nu\rangle \\ |\mu'\rangle|\nu'\rangle}} K_{\mu\nu\mu'\nu'}(t) ([\hat{H}_I(t), \cdot])_{\mu\nu\mu'\nu'} \quad (43)$$

The characteristic functional can now be written

$$\Phi[K(t)] \equiv \sum_{|r\rangle} \langle r | \mathbb{T} \exp \left\{ \frac{1}{\hbar} \int_0^\infty ds L(s) \right\} \sigma_R^{eq} | r \rangle \quad (44)$$

It is necessary to use the t -ordered exponential here because L is still a superoperator in the remainder Hilbert space spanned by the eigenkets $|r\rangle$.

The characteristic functional, however, is an ordinary scalar function.

The first two moments of $[\hat{H}_I(t), \cdot]$, in the sense of (31), are given by

$$\langle ([\hat{H}_I(t), \cdot])_{\mu\nu\mu'\nu'} \rangle = \frac{\delta}{\delta K_{\mu\nu\mu'\nu'}(t)} \phi[K(t)] \Big|_{K=0} \quad (45)$$

$$\langle ([\hat{H}_I(t), \cdot])_{\mu\nu\theta\beta} ([\hat{H}_I(s), \cdot])_{\theta\beta\mu'\nu'} \rangle = \frac{\delta^2}{\delta K_{\mu\nu\theta\beta}(t) \delta K_{\theta\beta\mu'\nu'}(s)} \phi[K(t)] \Big|_{K=0}$$

in which it is indicated that K is set equal to 0 after the functional derivatives have been taken. It is straightforward to check that this procedure leads to (36) and (37) for the first two cumulants.

The chief advantage to using characteristic functionals accrues when they can be expressed in closed form, i.e., the summations in (44) can actually be performed. The most frequently encountered case where this is so is the Gaussian case⁵. In the Gaussian case, $\phi[K(t)]$ turns out to be an ordinary exponential of a quadratic form in K . This does not mean that the operator cumulants associated with $[\hat{H}_I(t), \cdot]$ vanish after 2nd order.⁵ In a non-operator context, a Gaussian characteristic functional leads to cumulants which do vanish after the 2nd cumulant, and this fact is often used to characterize such processes as Gaussian. When non-commutativity is present, as it is with operators, a Gaussian process will have non-vanishing higher order cumulants, but will have a Gaussian characteristic functional.^{5,8} There has been much confusion in the literature with respect to these facts.

Connected Time Graphs and Fermi's Golden Rule

In whatever context cumulants are used, they may possess an extremely useful property, called the "cluster" property or the "irreducible graph" property. In the present context, this property relates to the time dependence of the

cumulant operators. The n^{th} order cumulant in (33) can be written in terms of an $n-1$ fold integral over an integrand which is a function of n time variables. The "cluster" property⁷, or "connected time graph" property, states that the integrand asymptotically vanishes whenever the n time variables are not clustered together within an interval of order τ_c , the correlation time.

Looking back at $G^{(2)}$ in (37), the connected time graph property implies that for $t \gg \tau_c$, $G^{(2)}$ is independent of t . This is because the s -integration only contributes a non-vanishing value for $|t-s| \leq \tau_c$. For $|t-s| \gg \tau_c$, the integrand virtually vanishes. As a result, for $t \gg \tau_c$, $G^{(2)}(t)$ is essentially a constant. The higher order cumulants, $G^{(n)}$ with $n > 2$, exhibit this property in a similar way, except that generally $G^{(n)}$ is a constant times a factor of order $\hat{\tau}_c^{n-2}$, whenever $t \gg \tau_c$. $\hat{\tau}_c$ is a dimensionless quantity containing the factor τ_c . The other factors are the ratio of a matrix element of \hat{H}_I and h , i.e.,

$$\hat{\tau}_c = \frac{E}{h} \tau_c \quad (46)$$

in which E is a measure of the extent of the range of values produced by matrix elements of \hat{H}_I . We can be more explicit with particular models.

The preceding points have the following significance. If τ_c is so short compared with all other time scales in a particular problem, that it can be approximately set equal to zero, then we can view the result as a Markov process.⁵ When this is true, the cumulants of order greater than $n = 2$ vanish because they are proportional to $\hat{\tau}_c^{n-2} \sim 0$. Thus, in the Markov approximation, $G^{(2)}$ alone provides the full result of the cumulant expansion. In examples in which the interaction between the subsystem and the remainder of the complete

system is electromagnetic, the coupling strength can be measured by the fine structure constant, α . The cumulant $G^{(2m)}$ works out to be of order α^m . Usually, we argue that this makes the higher order terms small because $\alpha \sim \frac{1}{137}$. However, we see here that $G^{(2m)}$ is really of order $\alpha^m \hat{\tau}_c^{2(m-1)}$ and that α alone does not tell the whole story. In fact, if τ_c were a sufficiently long time, then it could easily be more important than α , resulting in important contributions to the cumulant expansion by cumulants of high order. Generally, we find that α and $\hat{\tau}_c$ both conspire to render higher order cumulants negligible, when $\hat{\tau}_c$ is small, making the validity of the $G^{(2)}$ approximation even more secure.

We might have tried to solve (29) for the reduced density matrix using ordinary perturbation theory. Fermi's Golden Rule could have been applied to get the rate of change to lowest orders in the fine structure constant. It is well known that in many situations, this procedure gives good results. The cumulant approach leads to $G^{(2)}$ to lowest order in the fine structure constant. $G^{(2)}$ agrees with Fermi's Golden Rule, but it does so while it is the argument of an evolution exponential, as in (32). If this exponential is expanded, the linear term in $G^{(2)}$ is precisely Fermi's Golden Rule for the perturbation expansion. However, all the higher order iterations of the expansion of the exponential of $G^{(2)}$ yield an infinite subseries of terms from the ordinary perturbation expansion. In this sense, $G^{(2)}$, is a much stronger result than Fermi's Golden Rule, and entirely justifies using Fermi's Golden Rule as the argument of an exponentiated rate law, as is so often done. Moreover, for small $\hat{\tau}_c$, we see that $G^{(2)}$ is a very good approximation to the entire cumulant expansion. This is why Fermi's Golden Rule works so well in so many cases.

In the model calculations which follow, it will be necessary to establish the size of $\hat{\tau}_c$ in order to apply these results.

The Models

The model calculations which follow are contractions of exact dynamical equations. First, the Hamiltonian for the complete system is given for each model. Next, the specific nature of the contraction for each model is stipulated. Finally, the characteristic functional for each model is constructed. The characteristic functional can be used to obtain transition rates and energy shifts to any order in the coupling constants by application of functional differentiation. Expressions for these kinds of quantities are exhibited for these simple models.

All apparent stochastic properties are the result of the contraction of the description. In each case this contraction is accomplished by construction of reduced density matrices. No a priori stochasticity is inserted into the original, complete system Hamiltonians. For this reason, these calculations explain the origin of stochasticity, and irreversibility, in quantum statistical mechanics, and show how these properties emerge from the underlying exact, reversible dynamics.

Two models are considered. In each model the subsystem is a simple quantum mechanical system while the remainder of the complete system is a quantized boson field comprised of either photons or phonons. The boson field serves as a thermal reservoir. The reduced density matrices to be constructed describe the simple subsystem behavior. It is shown that the contractions employed, which eliminate the boson fields, result in Gaussian, non-Markovian processes for the subsystem description.

The total Hamiltonians for the complete system in the two models ^{are} space:

Model 1⁵⁾

$$H = H_B \otimes 1_P + 1_B \otimes H_P + H_I \quad (47)$$

in which H_B is the Hamiltonian for the interaction of a spin $\frac{1}{2}$ magnetic moment with an external magnetic field $\vec{B} = B\hat{k}$,

$$H_B = \frac{e\hbar}{2mc} \vec{\sigma} \cdot \vec{B} = \frac{e\hbar B}{2mc} \sigma_z \quad (48)$$

where σ_z is a Pauli matrix. H_P is the Hamiltonian for photons

$$H_P = \sum_{\vec{k}\lambda} \hbar \omega_k \left(a_{\vec{k}\lambda}^\dagger a_{\vec{k}\lambda} + \frac{1}{2} \right) \quad (49)$$

in which $a_{\vec{k}\lambda}^\dagger$ is the creation operator for a photon of polarization λ , propagation vector \vec{k} , and frequency $\omega_k = c|\vec{k}|$. The interaction Hamiltonian, H_I , is

$$H_I = \frac{e\hbar}{2mc} \vec{\sigma} \cdot \vec{B}(\vec{r}) \quad (50)$$

in which $\vec{B}(\vec{r})$ is the magnetic field associated with the photons through the plane wave expansion in volume V .

$$\vec{B}(\vec{r}) = \sqrt{\frac{2\pi\hbar}{V}} \sum_{\vec{k}\lambda} \frac{1}{\sqrt{\omega_k}} c \{ i \vec{k} \times \hat{\epsilon}_{\vec{k}\lambda} \exp[i \vec{k} \cdot \vec{r}] a_{\vec{k}\lambda} - i \vec{k} \times \hat{\epsilon}_{\vec{k}\lambda} \exp[-i \vec{k} \cdot \vec{r}] a_{\vec{k}\lambda}^\dagger \} \quad (51)$$

in which $\hat{\epsilon}_{\vec{k}\lambda}$ is the polarization unit vector for polarization λ and propagation vector \vec{k} . The interaction Hamiltonian in the interaction picture, in accord with (12), is

$$\hat{H}_I(t) = \frac{e\hbar}{2mc} e^{i\Omega t \sigma_z} \vec{\sigma} e^{-i\Omega t \sigma_z} \cdot \sqrt{\frac{2\pi\hbar}{V}} \sum_{\vec{k}\lambda} \frac{c}{\sqrt{\omega_k}} i(\vec{k} \times \hat{\epsilon}_{\vec{k}\lambda}) \left(a_{\vec{k}\lambda} e^{i\vec{k} \cdot \vec{r} - i\omega_k t} - a_{\vec{k}\lambda}^\dagger e^{-i\vec{k} \cdot \vec{r} + i\omega_k t} \right) \quad (52)$$

where $\Omega \equiv \frac{eB}{2mc}$.

Model 2⁵⁾

$$H = H_C \otimes 1_P + 1_C \otimes H_P + H_I \quad (53)$$

in which H_C is the Hamiltonian for a charged particle with charge Ze in an external potential $\phi(\vec{r})$,

$$H_C = \frac{\vec{p} \cdot \vec{p}}{2m} + Ze \phi(\vec{r}) \quad (54)$$

H_P is the Hamiltonian for phonons

$$H_P = \sum_{\vec{k}} \hbar \omega_{\vec{k}} (b_{\vec{k}}^\dagger b_{\vec{k}} + \frac{1}{2}) \quad (55)$$

in which $b_{\vec{k}}^\dagger$ is the creation operator for a bulk phonon with wave vector \vec{k} and frequency $\omega_{\vec{k}} = c|\vec{k}|$. The speed of sound, c , in this relation is determined by the bulk modulus B and the average mass density ρ_0 through the identity $c^2 = B/\rho_0$. The interaction Hamiltonian is

$$H_I = Ze \int d^3r' \frac{\delta\rho(\vec{r}')}{|\vec{r}-\vec{r}'|} \quad (56)$$

in which

$$\delta\rho(\vec{r}') = -q_0 \frac{1}{\sqrt{\rho_0} c} \sum_{\vec{k}} \sqrt{\frac{\hbar}{2V\omega_{\vec{k}}}} \left(\omega_{\vec{k}} e^{i\vec{k} \cdot \vec{r}'} b_{\vec{k}} + \omega_{\vec{k}} e^{-i\vec{k} \cdot \vec{r}'} b_{\vec{k}}^\dagger \right) \quad (57)$$

and q_0 is the average charge density of the medium. The summation over \vec{k} is bounded by the Debye frequency, i.e., $\omega_{\vec{k}} \leq \omega_D$. The interaction in (56) is the instantaneous Coulomb interaction between the bulk, longitudinal phonons and the charged particle. The interaction Hamiltonian, in the interaction picture, in accord with (12), is

$$\begin{aligned} \hat{H}_I(t) = & -Ze q_0 \frac{1}{\sqrt{\rho_0} c} \sum_{\vec{k}} \sqrt{\frac{\hbar \omega_{\vec{k}}}{2V}} \theta(\omega_D - \omega_{\vec{k}}) \frac{4\pi}{k^2} \left(\exp\left[\frac{i}{\hbar} t H_C\right] e^{i\vec{k} \cdot \vec{r}} \exp\left[\frac{i}{\hbar} t H_C\right] e^{-i\omega_{\vec{k}} t} b_{\vec{k}} \right. \\ & \left. + \exp\left[\frac{i}{\hbar} t H_C\right] e^{-i\vec{k} \cdot \vec{r}} \exp\left[-\frac{i}{\hbar} t H_C\right] e^{i\omega_{\vec{k}} t} b_{\vec{k}}^\dagger \right) \quad (58) \end{aligned}$$

in which $\theta(\omega_D - \omega_{\vec{k}})$ is the Heaviside step function which guarantees that $\omega_{\vec{k}} \leq \omega_D$. In model 1, a spin $\frac{1}{2}$ magnetic moment is interacting with the magnetic fluctuations of the photon field. These fluctuations give rise to an anomalous magnetic moment which ~~will~~ ^{may} be calculated. The treatment here is non-relativistic, which is known to give the wrong value for the anomalous magnetic moment, even the wrong sign. A relativistic treatment is known to give agreement with measurements. Our purpose here, however, is to exhibit the technique described in this paper, and this can be most simply achieved by looking at the details in Model 1. For our choice for the analogue of $\sigma_R^{eq.}$ in (23), we choose the pure state density matrix generated by the vacuum state of the photon field

$$\sigma_R^{eq.} \equiv |0\rangle \langle 0| \quad (59)$$

in which $|0\rangle$ denotes the Dirac vacuum state for the quantized photon field. Thus, the magnetic fluctuations are vacuum fluctuations which effect the magnetic moment.

In Model 2, the phonons are filling the role of a thermal reservoir. For the analogue of $\sigma_R^{eq.}$ in (23), we choose the canonical density matrix for phonons at temperature T .

$$\sigma_R^{eq.} = \prod_{\vec{k}} \exp[-\beta \hbar \omega_{\vec{k}} b_{\vec{k}}^\dagger b_{\vec{k}}] (1 - \exp[-\beta \hbar \omega_{\vec{k}}]) \quad \beta \equiv \frac{1}{k_B T} \quad (60)$$

When the temperature, T , is set equal to zero, the expression in (60) reduces to the phonon analogue of the vacuum density matrix for photons given in (59). Because of this, we can deduce results for Model 1 at non-zero temperatures by comparison with the results for Model 2.

The characteristic functionals for these two models are constructed

according to the procedure⁵ exhibited in equations (42), (43), and (44).

For Model 1, define the subsystem matrices

$$M_{\mu\mu'}^{\vec{k},\lambda}(t) = -i \frac{eh}{2m} \sqrt{\frac{2\pi\hbar}{V\omega_{\vec{k}}}} (\vec{k} \times \hat{\epsilon}_{\vec{k}}) \cdot \left(e^{i\Omega t \sigma_z} \vec{\sigma} e^{-i\Omega t \sigma_z} \right)_{\mu\mu'} \exp[-i(\vec{k} \cdot \vec{r} - \omega_{\vec{k}} t)] \quad (61)$$

$$N_{\mu\mu'}^{\vec{k},\lambda}(t) = i \frac{eh}{2m} \sqrt{\frac{2\pi\hbar}{V\omega_{\vec{k}}}} (\vec{k} \times \hat{\epsilon}_{\vec{k}}) \cdot \left(e^{i\Omega t \sigma_z} \vec{\sigma} e^{-i\Omega t \sigma_z} \right)_{\mu\mu'} \exp[i(\vec{k} \cdot \vec{r} - \omega_{\vec{k}} t)]$$

In addition, define

$$K_M^f(\vec{k},\lambda)(t) = \frac{1}{\hbar} K_{\mu\nu\mu'\nu'}(t) \delta_{\nu\nu'} M_{\mu\mu'}^{\vec{k},\lambda}(t) \quad (62)$$

$$K_M^b(\vec{k},\lambda)(t) = \frac{1}{\hbar} K_{\mu\nu\mu'\nu'}(t) \delta_{\mu\mu'} M_{\nu'\nu}^{\vec{k},\lambda}(t)$$

$$K_N^f(\vec{k},\lambda)(t) = \frac{1}{\hbar} K_{\mu\nu\mu'\nu'}(t) \delta_{\nu\nu'} N_{\mu\mu'}^{\vec{k},\lambda}(t)$$

$$K_N^b(\vec{k},\lambda)(t) = \frac{1}{\hbar} K_{\mu\nu\mu'\nu'}(t) \delta_{\mu\mu'} N_{\nu'\nu}^{\vec{k},\lambda}(t)$$

These definitions greatly simplify the analysis based on (43). The superscripts f and b refer to forward and backward time ordering which result from the time ordering in (44) and the commutator in (43) in close parallel with identity (20). For the characteristic functional for Model 1, we obtain⁵⁾

$$\begin{aligned} \phi[K(t)] = \exp[- \int_0^\infty dt_1 \int_0^{t_1} dt_2 \sum_{\vec{k},\lambda} [& K_N^f(\vec{k},\lambda)(t_1) K_M^f(\vec{k},\lambda)(t_2) + K_M^b(\vec{k},\lambda)(t_1) K_N^b(\vec{k},\lambda)(t_2) \\ & - K_M^f(\vec{k},\lambda)(t_1) K_N^b(\vec{k},\lambda)(t_2) - K_N^b(\vec{k},\lambda)(t_1) K_M^f(\vec{k},\lambda)(t_2)]] \quad (63) \end{aligned}$$

which is manifestly Gaussian.

For Model 2, define the subsystem matrices

$$M_{\mu\mu'}^{\vec{k}}(\vec{r},t) = -Ze q_0 \frac{1}{\sqrt{\rho_0} c} \sqrt{\frac{\hbar\omega_{\vec{k}}}{2V}} \frac{4\pi}{k^2} (\exp[-i\vec{k} \cdot \vec{r}])_{\mu\mu'} \exp[i(\omega_{\mu\mu'} + \omega_{\vec{k}})t] \quad (64)$$

$$N_{\mu\mu'}^{\vec{k}}(\vec{r}, t) = -Ze q_0 \frac{1}{\sqrt{\rho_0} c} \sqrt{\frac{\hbar\omega_{\vec{k}}}{2V}} \frac{4\pi}{k^2} (\exp[i\vec{k}\cdot\vec{r}])_{\mu\mu'} \exp[i(\omega_{\mu\mu'} - \omega_{\vec{k}})t]$$

in which $\omega_{\mu\mu'} = \frac{1}{\hbar} (E_{\mu} - E_{\mu'})$ where E_{μ} and $E_{\mu'}$ are eigenvalues of H_C corresponding with the eigenkets $|\mu\rangle$ and $|\mu'\rangle$ of H_C . We will also need the analogue of (62) for this model. The corresponding definitions will look just like (62) except that it will not be necessary to include the parameter λ . For the characteristic functional for Model 2 we obtain⁵⁾

$$\begin{aligned} \Phi[K(t)] = \exp \left[- \int_0^{\infty} dt_1 \int_0^{t_1} dt_2 \sum_{\vec{k}} \theta(\omega_D - \omega_{\vec{k}}) \left[\left(K_N^f(\vec{k})(t_1) - K_N^b(\vec{k})(t_1) \right) \left(K_M^f(\vec{k})(t_2) \frac{1}{1-\rho(\vec{k})} - K_M^b(\vec{k})(t_2) \frac{\rho(\vec{k})}{1-\rho(\vec{k})} \right) \right. \right. \\ \left. \left. + \left(K_M^b(\vec{k})(t_1) - K_M^f(\vec{k})(t_1) \right) \left(K_N^b(\vec{k})(t_2) \frac{1}{1-\rho(\vec{k})} - K_N^f(\vec{k})(t_2) \frac{\rho(\vec{k})}{1-\rho(\vec{k})} \right) \right] \right] \quad (65) \end{aligned}$$

in which $\rho(\vec{k}) \equiv \exp[-\beta\hbar\omega_{\vec{k}}]$. This characteristic functional is also Gaussian.

The structure of the $\Phi[K]$ in (65) reduces to the structure of the $\Phi[K]$ in (63) when $T \rightarrow 0$ ($\beta \rightarrow \infty$). Conversely, the results in Model 1 for $T \neq 0$, can be obtained by generalizing (63) to (65), provided that the $\theta(\omega_D - \omega_{\vec{k}})$ is left out in the photon case.

In both (63) and (65), the \vec{k} -summation may be replaced by an integral in the infinite volume limit according to the limit

$$\sum_{\vec{k}} \dots \rightarrow \frac{V}{(2\pi c)^3} \int_0^{\omega_M} d\omega \omega^2 \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \dots \quad (66)$$

In Model 1; the upper limit for the ω -integration, ω_M , is set equal to $\frac{mc^2}{\hbar}$ as is customary in non-relativistic treatments of this kind. In Model 2, ω_M is simply ω_D , the Debye frequency. The c in (66) is the speed of light for Model 1, and is the speed of sound for Model 2. The remaining factor of V which is explicit in (66), cancels the normalization factors of \sqrt{V} in (61)

or (64) when these expressions are put into (63) and (65) by means of (62).

We may view the results of this treatment of the \vec{k} -summations in (63) and (65) as time correlation functions. For example, the \vec{k} -summation of the product $K_N^{f(\vec{k},\lambda)}(t_1) K_M^{f(\vec{k},\lambda)}(t_2)$ in (63) contains the ω -integral:

$$\int_0^{\omega_M} d\omega \omega^3 e^{-i\omega(t_1-t_2)} = \quad (67)$$

$$\frac{\omega_M^3}{-i(t_1-t_2)} e^{-i\omega_M(t_1-t_2)} - 3 \frac{\omega_M^2}{(-i(t_1-t_2))^2} e^{-i\omega_M(t_1-t_2)} + 6 \frac{\omega_M}{(-i(t_1-t_2))^3} e^{-i\omega_M(t_1-t_2)}$$

$$- \frac{6}{(-i(t_1-t_2))^4} e^{-i\omega_M(t_1-t_2)} + \frac{6}{(-i(t_1-t_2))^4}$$

An application of the ~~Riemann~~ Riemann-Lebesgue lemma to the left-hand side of (67) guarantees that the integral vanishes in the limit $t_1-t_2 \rightarrow \infty$. This is manifestly the case on the right-hand side of (67). Moreover, the rapid oscillations of the numerator exponentials on the right-hand side of (67) causes the correlation to die off much more rapidly than the simple power law denominators alone would suggest. These features are characteristic of the correlation functions for both models, and justify the existence of τ_c in (39).

Typical Model Results

The reduced density matrix for Model 1 is a 2x2 density matrix. Since it is Hermitean, it may be expanded in terms of the Pauli matrices. When this is done, the coefficients of the Cartesian Pauli matrices turn out to be the Cartesian components of the magnetization vector for the spin $\frac{1}{2}$ magnetic moment.⁵⁾ Working out the 2nd cumulant contribution to (29) yields the Bloch magnetization equations:

$$\frac{d}{dt} M_z(t) = -\frac{1}{T_R} M_z(t) + \frac{1}{T_R} \quad (68)$$

$$\frac{d}{dt} M_x(t) = \frac{1}{\hbar} (E_2 - E_1) M_y(t) + \frac{1}{\hbar} (\Delta E_2 - \Delta E_1) M_y(t) - \frac{1}{T_r} M_x(t) \quad (69)$$

$$\frac{d}{dt} M_y(t) = -\frac{1}{\hbar} (E_2 - E_1) M_x(t) - \frac{1}{\hbar} (\Delta E_2 - \Delta E_1) M_x(t) - \frac{1}{T_r} M_y(t) \quad (70)$$

$$\text{in which } E_2 = \frac{e\hbar}{2mc} \circlearrowleft B, \quad E_1 = -\frac{e\hbar}{2mc} \circlearrowleft B, \quad \frac{1}{T_r} = \frac{2}{3} \frac{e^5 \hbar}{m^5 c^8} B^3 \quad \text{and} \quad (71)$$

$$\Delta E_2 = \frac{1}{3\pi} \frac{e^2 \hbar^3}{c^5 m^2} \int_0^{\omega_M} \omega^3 \frac{1}{E_2 - E_1 - \hbar\omega} d\omega$$

$$\Delta E_1 = \frac{1}{3\pi} \frac{e^2 \hbar^3}{c^5 m^2} \int_0^{\omega_M} \omega^3 \frac{1}{E_1 - E_2 - \hbar\omega} d\omega$$

The "longitudinal" term, M_z comes from the diagonal part of the reduced density matrix, whereas the "transverse" terms, M_x and M_y , come from the off-diagonal part of the reduced density matrix. The full equilibrium result of these equations is $M_z = 1$ and $M_x = M_y = 0$. This is a zero temperature result. For $T \neq 0$, Model 2 leads to a related result for full equilibrium, a reduced density matrix identical with the canonical density matrix in terms of the eigenenergies of H_C . The energy shift formulas in (71) provide corrections of first order in the fine structure constant, α , to the eigenenergies of H_B . These corrections do not show up in the equilibrium result, but do contribute an "anomalous magnetic moment" correction⁵⁾ to any resonance experiment. Similarly, for Model 2 with $T \neq 0$, any resonance measurement would involve order α corrections to the energy difference factors which are the analogues to $\Delta E_2 - \Delta E_1$ in (69) and (70). In general, $G^{(2n)}$ provides equilibrium results with energies corrected to order α^{n-1} and resonance results with energy differences corrected to order α^n . Equations (68) - (70) are the most simple example of this behavior. However, we neglect higher order cumulants, $G^{(2n)}$ for $n > 1$, not so much because they lead to corrections of order α^{n-1} and α^n , but because the correlation time built into (63) and (65) is so short compared with the relaxation time, which

is $\frac{1}{T_r}$ in (68) - (70). Obviously, by taking B very small, the relaxation time, T_r , becomes very long, according to (71). The correlation time, on the other hand is determined by $\omega_M = \frac{mc^2}{\hbar}$. The magnitude of this correlation time is found by exploring the details of the correlation functions discussed in the last section. For Model 1, we have⁵⁾

$$\int_0^{\omega_M} d\omega \omega^3 \cos[(\omega + \Delta\omega)\tau] = C(\tau) \quad (72)$$

for purely "longitudinal" couplings in (29). These are the diagonal density matrix element transition rate correlation functions. The term $\Delta\omega$ is equal to $\frac{1}{\hbar} (E_2 - E_1) = \frac{eB}{mc}$. The integral is elementary and gives the analogue of the real part of (67):

$$C(\tau) = \frac{\omega_M^3}{\tau} \sin[(\omega_M + \Delta\omega)\tau] + \frac{3\omega_M^2}{\tau^2} \cos[(\omega_M + \Delta\omega)\tau] - \frac{6\omega_M}{\tau^3} \sin[(\omega_M + \Delta\omega)\tau] \quad (73)$$

$$- \frac{6}{\tau^4} \cos[(\omega_M + \Delta\omega)\tau] + \frac{6}{\tau^4} \cos[(+\Delta\omega)\tau]$$

It can be shown that only $-\Delta\omega$ leads to interesting results, whereas $+\Delta\omega$ leads to results which can not satisfy energy conservation. For short times, all of the terms in (73) must be combined to yield finite behavior, but for long times, the leading term dominates. An application of the Dirichlet integral lemma which is a companion of the Riemann-Lebesgue lemma implies

$$\frac{2}{\pi} \int_0^{\infty} d\tau g(\tau) \frac{\sin[(\omega_M + \Delta\omega)\tau]}{\tau} = g(0+) \quad (74)$$

for any smooth function $g(\tau)$. The scale of this approximation is set by

$\frac{1}{\omega_M} = \frac{\hbar}{mc^2}$. For an electron, we get $\omega_M^{-1} = 1.3 \times 10^{-21}$ sec, which is much less than T_r , even for very large magnetic fields, B.

The Markov limit of $G^{(2)}$ provides an irreversible dynamical description.

The non-zero temperature generalization of Model 1 which is analogous to Model 2 in structure can be used to construct a $G^{(2)}$ for $T \neq 0$. The corresponding reduced density matrix is irreversibly driven by $G^{(2)}$ toward the Helmholtz free energy minimum appropriate at temperature T . The equilibrium value agrees with the Boltzmann distribution in unshifted energies. Resonance values induced by external couplings involve energies shifted to first order in the fine structure constant α . The approach to equilibrium is accompanied by a monotropic decrease in the free energy.⁹ In an earlier, cruder model construction,⁹ I was able to prove this irreversible, monotone behavior of the Helmholtz free energy while commemorating Mark Kac's 60th birthday. It is a double pleasure to now commemorate his 65th.

References

1. Mark Kac, Probability and Related Topics in Physical Sciences, (Interscience Pub., London, 1959).
2. G. E. Uhlenbeck and G. W. Ford, Lectures in Statistical Mechanics, (Amer. Math. Soc., Providence, R. I., 1963). Chapter 1.
3. Mark Kac, Statistical Independence in Probability, Analysis, and Number Theory, (Amer. Math. Soc., 1959). Chapter 5.
4. R. C. Tolman, The Principles of Statistical Mechanics, (Oxford, London, 1938). Chapter IX.
5. R. F. Fox, Phys. Rev. 48, 179 (1978), Sections II.3 and II.7.
6. R. Kubo, J. Math. Phys. 4, 174 (1963).
7. R. F. Fox, J. Math. Phys. 17, 1148 (1976).
8. R. F. Fox, J. Math. Phys. 20, 2467 (1979).
9. R. F. Fox, J. Math. Phys. 16, 1214 (1976).

PLEASE READ INSTRUCTIONS ON REVERSE BEFORE COMPLETING

PART I—PROJECT IDENTIFICATION INFORMATION

1. Institution and Address Georgia Institute of Technology Atlanta, Georgia 30332	2. NSF Program Theoretical Physics	3. NSF Award Number PHY 7921541
	4. Award Period From 1/15/80 To 6/30/83	5. Cumulative Award Amount \$43,678
6. Project Title Fluctuation Phenomena in Thermal Physics		

PART II—SUMMARY OF COMPLETED PROJECT (FOR PUBLIC USE)

During the period of this grant, 10 papers were published and 3 Ph.D. theses were produced. Some of the work concerned hydrodynamic fluctuations away from equilibrium and light scattering measurements for such processes. Other work dealt with coupled rotational and translational diffusion of molecules in liquids. The phenomenon of long time tails was studied and an analysis of computer simulations, theory, and experiments related to this phenomenon was published. Mathematical methods for the study of multiplicative stochastic processes were developed. These processes arise in the study of laser noise problems for which a new theory has been developed. The attached report contains greater detail about each of these topics.

PART III—TECHNICAL INFORMATION (FOR PROGRAM MANAGEMENT USES)

1. ITEM (Check appropriate blocks)	NONE	ATTACHED	PREVIOUSLY FURNISHED	TO BE FURNISHED SEPARATELY TO PROGRAM	
				Check (✓)	Approx. Date
a. Abstracts of Theses		X 3			
b. Publication Citations				✓ 10	Oct. 1983
c. Data on Scientific Collaborators					
d. Information on Inventions					
e. Technical Description of Project and Results		X			
f. Other (specify)					
2. Principal Investigator/Project Director Name (Typed) Ronald F. Fox		3. Principal Investigator/Project Director Signature		4. Date Oct. 18, 1983	

Annual Report for NSF Grant PHY-7921541 by Ronald F. Fox

"Fluctuation Phenomena in Thermal Physics"

This report covers the third and final year of activity under this grant.

In December, 1982, a graduate student, Barry J. Cown, concluded his Ph.D. dissertation: "Stochastic Near-Field Theory and Techniques for Wideband Electromagnetic Emitters of In-band and Out-of-band Frequencies", under my supervision. Dr. Cown is an applied physicist working on radar problems at the Engineering Experiment Station here at Georgia Tech. This work represents a very applied utilization of my expertise in stochastic processes.

Another graduate student, Byron Burel, who will earn his Ph.D. in June, 1983, has concluded his studies of light scattering from a fluid in a thermal gradient. These studies were designed to test a theory for hydrodynamic fluctuations in a fluid far from thermal equilibrium. The fluid is driven to a steady state by imposition of a thermal gradient. The thermal gradient induces spatial dependence in all fluid properties, including the dissipative parameters such as viscosity and heat conductivity. Cohen, Dorfman, and Kirkpatrick have used kinetic theory to show that a major enhancement of Rayleigh scattering is to be expected under appropriate conditions. This is to be contrasted with the earlier work which showed the modification of the Brillouin side bands under the same conditions, but to a much less detectable level. I reviewed the Brillouin situation in the paper: "Testing Theories of Nonequilibrium Processes with Light-Scattering Techniques", J. Phys. Chem. 86, 2812 (1982). Recently, Ronis and Procaccia confirmed the work of Cohen, Dorfman and Kirkpatrick for the Rayleigh line. Both groups, however, have chosen a situation in which the spatial dependence of the transport coefficients can be ignored. Burel and I have attacked the much harder problem where this

dependence is not ignorable.

The problem of hydrodynamic fluctuations away from equilibrium in the non-linear regime continues to provide debate among theorists. My paper: "Stress-Strain Fluctuations in Non-linear Hydrodynamics", *Physica* 112A, 505 (1982) discusses the approach of Bedeaux and Mazur which I have criticized.

Having decided that hydrodynamics may not be the best context in which to test non-equilibrium fluctuation theory, work commenced on dye laser fluctuations. This context simultaneously involves non-equilibrium fluctuations, multiplicative stochastic processes, and non-Markovian processes. In the last regard, a paper: "Correlation Time Expansion for Non-Markovian, Gaussian, Stochastic Processes", has been accepted for publication by *Physics Letters*. Collaboration with Raj Roy, an experimentalist who recently joined our School of Physics provides added impetus to this work on laser noise.

Two papers appeared from Ph.D. work of Ulrich Steiger, who graduated two years ago. They were: "Coupled Translation and Rotational Diffusion in Liquids", *J. Math. Phys.* 23, 296 (1982), and "Boson Operator Representation of Brownian Motion", *J. Math Phys.* 23, 1678 (1982). This work provides a remarkable representation of diffusion processes in quantum field theory language. The objective of these studies is a better understanding of diffusion controlled reactions in liquid solvents.

My program of study of correlation functions using cumulant methods lead to a direct confrontation of the question of long-time tails. In June, 1982, I attended a conference on Nonlinear Phenomena in Fluids at Boulder, Colorado and gave a talk: "The Long Time Tail Conundrum in Nonequilibrium Statistical Mechanics", to appear in *Physica A*. I addressed the computer simulation,

theoretical, and experimental bases for long time tails in diffusion processes. I enunciated a variety of qualms regarding the firmness of each type of evidence. A much more detailed account of my objections: "Long Time Tails and Diffusion", will soon appear in Phys. Rev. A. Three incidental papers were also produced during this period of research. "The Ideal Gas and the Second Law of Thermodynamics", Am. J. Phys. 50, 804 (1982), presented my thoughts after my encounter with I. Prigogine in 1981. "Elementary Analysis of Electric Dipole Transitions Induced by Semi-Classical $\vec{A}\cdot\vec{p}$ and $\vec{E}\cdot\vec{r}$ Perturbations", written with Bill Harter, is submitted to Am. J. Phys. It clears up a long standing, published confusion about $\vec{A}\cdot\vec{p}$ versus $\vec{E}\cdot\vec{r}$. The resolution involves a gauge transformation, originally devised by Maria Mayer in 1931! This paper is relevant in multiphoton processes in molecules. The third paper is a remark about a paper recently published in Phys. Rev. Letters. My paper, with H. A. Gersch: "Critique of 'Quantum Statistics for Distinguishable Particles'", shows that the paper critiqued does not achieve its claimed objective, the establishment of quantum statistics using an argument based on distinguishable particles, but instead contains a blunder.

Overall the year has proved productive. Research activity has shifted towards quantum noise problems.

"Analysis of Coupled Translational and Rotational Diffusion Using Operator Calculus," by Ulrich R. Steiger, Ph.D. 1981.

SUMMARY

The equations for coupled translational and rotational diffusion of asymmetric molecules immersed in a fluid are obtained. The method used begins with the Kramers-Liouville equation and leads to the generalized Smoluchowski equation for diffusion in the presence of potentials. Both external potentials and intermolecular potentials are considered. The contraction of the description from the Kramers-Liouville equation to the Smoluchowski equation is achieved by using a combination of operator calculus and cumulants. Explicit solutions of these equations are obtained in the two-dimensional case. The formalism also allows the calculation of corrections to the generalized Smoluchowski equation. Smoluchowski's result is precisely the second cumulant, in the cumulant expansion.

The next non-vanishing term, the fourth cumulant, leads to diffusion equations with position dependent diffusion coefficients. The higher order cumulants lead to evolution equations for the reduced probability density which contain partial derivatives of order m with $m \geq 3$. Explicit expressions are given up to the sixth order in the cumulant expansion for translation diffusion. From a practical point of view, this formalism is very useful because partial differential equations can be solved numerically by using a finite element calculation.

The contraction of the Liouville-Kramers description into the Smoluchowski description is achieved by using a creation-destruction operator representation. In an appropriately defined inner product space, these operators possess the operator algebra of boson operators in quantum field theory. The discovery of this representation for the Liouville-Kramers description including rotational effects greatly facilitated obtaining the Smoluchowski contraction

"Stochastic Near-Field Theory and Techniques for Wideband Electromagnetic Emitters at In-Band and Out-of-Band Frequencies," by Barry J. Cown, Ph.D. 1982.

SUMMARY

Antenna systems play an important role in the defense of both the civilian and military populations against incoming threat missiles. Typical antenna installations aboard surface ships or at Army field sites may have various radar, navigation, and communication antenna systems located within a relatively small area. The scattering and coupling of electromagnetic energy among the various antennas can cause severe degradations in the ability of the antenna systems to perform their designated functions.

While all of the antenna systems are important, the directive microwave radar antenna systems are crucial. These systems have the tasks of detecting and tracking incoming missiles and, in some installations, they also guide the defensive surface-to-air missiles to the incoming threat missiles. Reliable detection and accurate tracking of a target can be hindered, or even prevented altogether, by the coupling of unwanted signals among antennas operating in the same frequency band (in-band coupling) or different frequency band (in-band to out-of-band coupling or out-of-band to out-of-band coupling). Further, the electromagnetic energy scattered by the antennas, the support structures, and other scattering objects can result in reduced detection range, tracking errors, and elevated sidelobe radiation. The increased sidelobe radiation can render the installation more susceptible to enemy anti-radiation missiles (ARM) which depend on high average sidelobe radiation levels for detection of and guidance to their target.

The foregoing considerations make it imperative that the radar antenna systems function in a reliable, predictable, and electromagnetically compatible manner (compatible implying the absence of "too much" coupling) if they are to be electromagnetically effective against threat missiles. Hence, electromagnetic compatibility (EMC) and electromagnetic effectiveness (EME) are issues of vital concern. Of course, the analysis and design (or re-design) of real-world antenna installations to achieve optimum EMC and EME is a very complex problem area in applied physics, even if one were to consider electromagnetic coupling and scattering at only a single in-band frequency. However, the modern trend is toward operation over a wide range of frequencies, i.e., toward wideband continuous wave (cw) or pulsed systems. In addition, the ability to predict and control the effects of radiation at

out-of-band frequencies has also become very important as the sensitivity and performance requirements of modern radar systems have increased.

An accurate electromagnetic coupling and scattering analysis of cosited antenna systems for a given operating in-band or out-of-band frequency requires, as the first step, a knowledge of the nominal, or "clear-site", antenna electric field pattern for each antenna that would be obtained in the absence of other antennas or scattering obstacles. The most accurate patterns are obtained via the planar near-field measurement technique whereby the antenna's electric field pattern is computed numerically from electric field data obtained on a plane located within the near-field, or "Fresnel zone", of the antenna. However, this technique was developed for measuring very stable, single mode antennas at a single operating frequency. Consequently, the application of the near-field technique to measure wideband in-band and out-of-band antennas should be investigated to identify and, if possible, to overcome special problems that can arise. For example, it is known from experiments and theoretical considerations that out-of-band antenna patterns can exhibit erratic, seemingly random behavior. Thus, it is anticipated that stochastic theory and analysis will be needed in order to characterize out-of-band antenna patterns properly and efficiently.

The second step in the analysis of coupling and scattering at a single frequency requires the use of the clear-site antenna patterns to compute coupling and scattering based on either the Plane Wave Spectrum (PWS) analysis technique or other established electromagnetic analysis techniques. Thus, the extension to wideband in-band and out-of-band frequencies needs to be studied. Again, a stochastic technique is indicated for analyzing in-band to out-of-band coupling, out-of-band to out-of-band coupling, or out-of-band scattering.

Advances in the state-of-the-art of wideband antenna analysis can be achieved through a basic theoretical and numerical research effort. Accordingly, theoretical and numerical analyses were performed to study the application of near-field theory and techniques to characterize the radiation and coupling characteristics of wideband, in-band and out-of-band pulsed or cw radiating systems. Specifically, stochastic theory and equations were developed for characterizing the radiation patterns of wideband cw or pulsed antennas over both in-band and out-of-band frequency intervals from measured data collected via near-field measurement techniques. The results are applicable to either phased array or reflector antennas. Three analytical techniques for analyzing the in-band and out-of-band coupling between pairs of cosited antennas

were studied. The three techniques are (1) the Plane Wave Spectrum (PWS), (2) the Spherical Wave Spectrum (SWS), and (3) Geometrical Theory of Diffraction (GTD). The existing theory and equations that are applicable to selected common waveguide components under normal in-band operation were extended to describe wideband out-of-band responses. Also, the theory and equations were formulated for computing the higher-order mode coefficients at the aperture of a waveguide radiating element from a knowledge of the measured far-field electric field of the radiating element when surrounded by a large conducting "ground" plane. Equations for describing the effects of near-field obstacles located in the antennas' forward half-plane on the performance of a wideband cw or pulsed antenna were derived via the Plane Wave Spectrum (PWS) analysis technique. The resultant statistical average pattern versus frequency is expressed explicitly in terms of the antenna system mode excitation statistical parameters.

"Fluctuations in a Nonequilibrium Steady State: Light Scattering from a Thermal Gradient," by Byron L. Burel, Ph.D. 1983.

SUMMARY

Near equilibrium fluctuation theory is extended into the nonequilibrium regime through the use of the postulate of local equilibrium, and the resulting nonequilibrium fluctuation theory is applied to a hydrodynamic system in which a steady state thermal gradient is present. This approach to nonequilibrium fluctuation theory is applied in an experimentally testable context to light scattering from water. In particular, the effects of the temperature dependence of the hydrodynamic transport coefficients upon the intensity of Rayleigh scattering are investigated.

This calculation is accomplished by using Fourier series to reduce the hydrodynamic equations to matrix form. The resulting matrix equation is solved for the mass density fluctuations, and the mass density autocorrelation function is then computed. The connection between the mass density autocorrelation function and the structure function describing the scattered light intensity is established, and the Rayleigh intensity is computed.

The result of this calculation indicates that the presence of a thermal gradient induces an enhancement of the Rayleigh intensity which may be a substantial fraction of the corresponding equilibrium intensity.