Testing Boltzmann's ergodic hypothesis with electron gas models

2006 J. Phys. A: Math. Gen. 39 4651

(http://iopscience.iop.org/0305-4470/39/17/S52)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 4651–4658 doi:10.1088/0305-4470/39/17/S52

Testing Boltzmann's ergodic hypothesis with electron gas models*

M Howard Lee

Physics Department, University of Georgia, Athens, GA 30602, USA[1]
and
Korea Institute for Advanced Study, Seoul 130-012, Korea

E-mail: mhlee@uga.edu

Received 30 August 2005, in final form 14 November 2005
Published 7 April 2006
Online at stacks.iop.org/JPhysA/39/4651

Abstract

Kubo proposed a physical approach to proving the validity of Boltzmann's ergodic hypothesis. It is to perform the time averages on dynamical functions, thereby avoiding the difficulties of measure theory inherent in the classical approach. To perform time averaging properly, one must have a general solution for the Heisenberg equation of motion such as by the recurrence relations method. A time averaging carried over with a recurrence relations solution is found to yield an ergodic condition in the form of an infinite product. It is linked to the energy transfer mechanisms, hence to the ergodicity itself. The electron gas models are fertile ground for testing ergodicity by this approach. For several static domains, we have evaluated the infinite product and drawn from them a general physical picture that underlies the ergodic hypothesis.

PACS numbers: 05.40.−a, 05.60.−k

1. Introduction

To be sure, Boltzmann's ergodic hypothesis is profound [1]. Evidently, he must have been the first to attempt to prove its validity. Seen today, more than a hundred years later, one is tempted to say that he should have approached it by performing time averages. Boltzmann instead suggested another idea. If successfully implemented, one could suppose that it would be equivalent to performing time averages. This turns out to be a very difficult task, a study which has come to be known as ergodic theory, then quasi ergodic theory, in mathematics [2]. The resulting studies perforce have become remote from the physical origin. It seems unlikely that ergodic or quasi ergodic theory could ever say whether the thermodynamic limit

* Dedicated to the memory of D ter Haar.

[1] Permanent address.

or irreversibility is a necessary *and* sufficient condition for ergodicity, or more specifically, what makes a model ergodic or how can a model be ergodic with respect to one variable but not with respect to another. The problem is, in Dorfman's words [3], 'still unsolved'.

So why not perform the time averages in the first place? Some years ago ter Haar [4] wrote that in doing so, 'one encounters serious difficulties'. Khintchin [5] seems to have been the first to suggest performing time averages over a broad class of correlation functions. Later Kubo [5] considered time averages more narrowly over response functions, that is, the dynamic susceptibilities. In linear response theory, these functions are related to inelastic scattering processes. Thus, by focusing on the time averages on them, one can gain the possibility of tying ergodic behaviour to energy transfer mechanisms, something that is natural to physical understanding.

By time averaging, Kubo deduced what has come to be known as Kubo's ergodic condition. That is, if the zero-frequency limit of a dynamic susceptibility is equal to its static counterpart, the system is ergodic. In the 1970s, there were a number of studies on magnetic models testing Kubo's condition but the results were ambiguous. Although his idea was sound, Kubo also did not have at his disposal a general solution to the Heisenberg equation. Thus, his performing time averaging was merely formal, not actually taken over the time evolution itself, thus, inevitably missing out on some essential factors as we shall see.

Let A be a dynamical variable in a many-particle system, defined by a Hamiltonian H. As is known, $A(t)$ the time evolution of A is obtained by solving the Heisenberg equation. If $A(t)$ were thus obtained for a given H, one could construct e.g. $\langle A^+(t)A\rangle$ where $^+$ denotes Hermitian conjugation and the brackets mean an ensemble average. Then by performing a time average, one could see whether it is equal to $\langle A^+A\rangle$. A straightforward operation is an unambiguous way of testing the hypothesis. But this path has not been taken—recall the words of ter Haar—because solving the Heisenberg equation has been a major challenge. But now this direct path is open since the recurrence relations method [6] can give a general solution for the Heisenberg equation.

Using this solution, we have recently carried out time averages on the response functions and established an ergodic condition [7]. Therewith, we are able to connect ergodic or nonergodic behaviour to the response of a system to an external perturbation. We also find that irreversibility is a necessary but not sufficient condition for ergodicity. The same may be said of the thermodynamic limit.

2. Ergodic hypothesis on response functions

We shall briefly summarize the response functions defined by linear response theory. Let us assume H is Hermitian. Let our system be perturbed by an external probe h, such that the interaction energy $V = hA$, where A is a dynamical variable of our system, through which the probe is coupled. The total energy is thus $H_{\text{tot}} = H + V$. If h is time independent, linear response theory gives the static response function χ_A. If h is time dependent, it gives the dynamic response function $\chi_A(t, t')$. If our system is *causal* and *stationary*, which we also assume, $\chi_A(t, t') = \chi_A(t - t')$, $t > t'$. In particular it has the form

$$\chi_A(t) = \begin{cases} \mathrm{i}/\beta\hbar\langle[A^+(t), A]\rangle & \text{if} \quad t > 0 \\ 0 & \text{if} \quad t \leqslant 0, \end{cases} \tag{1}$$

where $A(t) = \exp \mathrm{i}tH/\hbar A \exp -\mathrm{i}tH/\hbar$, $A = A(t = 0)$, $\beta = 1/kT$ and the angular brackets mean an ensemble average with respect to the states of H.

The ergodic hypothesis (EH) applied to the response functions is as follows:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \int_0^t \chi(t-t')\,\mathrm{d}t'\,\mathrm{d}t = \chi, \tag{2}$$

where the limit on T implies an irreversibility in the perturbed system after Boltzmann. By introducing a second function $R(t)$, with the properties $\mathrm{d}R(t)/\mathrm{d}t = -\chi(t)$ and $\mathrm{d}R(0)/\mathrm{d}t = 0$, the lhs of equation (2) can be expressed as

$$\tilde{\chi}(0) + R(T) - \tilde{R}(0)/T = \chi, \tag{3}$$

where $T \to \infty$ is implied on the lhs and e.g. $\tilde{\chi}(z) = L(z,t)\chi(t)$, L is the Laplace transform operator. We do not yet know whether equation (2) or its exact equivalent equation (3) is valid. To show that, we shall compare it with an exact expression. Let us add one more property to $R(t)$: $R(0) = \chi$. Assuming that such a function exists, we can arrive at the following:

$$\tilde{\chi}(0) + z\tilde{R}(z)|_{z=0} = \chi. \tag{4}$$

Equation (4) is exact since it follows from the definition of the function $R(t)$. If equations (3) and (4) are compared, clearly they do not agree. One must thus conclude that in general EH cannot be valid, at least, as applied to the response functions.

If however $\tilde{R}(0)$ is finite, the two equations can coincide, which gives an ergodic condition in the form

$$0 < \int_0^\infty R(t)\,\mathrm{d}t < \infty. \tag{5}$$

The lower bound arises from the fact that ergodicity is a consequence of $T \to \infty$, i.e., irreversibility. We shall later see that if $\tilde{R}(0) = 0$, there is a localization of energy. If $\tilde{R}(0) = \infty$, there is also no delocalization of energy. At both limits, EH fails. If our condition equation (5) is satisfied, it yields Kubo's condition $\tilde{\chi}(0) = \chi$. But Kubo's condition does not necessarily imply equation (5). We also note that irreversibility i.e. $R(t) \to 0$ as $t \to \infty$ is only a necessary but not sufficient condition for ergodicity.

What is $R(t)$? The function $R(t)$ which has the three key properties is the relaxation function defined as follows:

$$R_A(t) = (A(t), A), \tag{6}$$

where the inner product means the Kubo scalar product (KSP), a generalized form of fluctuations. In the theory of recurrence relations, the KSP realizes an inner product space. It is thus sufficient to know the time evolution of A on this space to determine $R(t)$. The recurrence relations method [6] gives such a solution, given in section 3. In recent years many have studied the relaxation function and its analogues like the memory function. We bring attention to a few of the more germane ones albeit incomplete [8–22].

3. The recurrence relations method

The techniques of the recurrence relations method are well described in several places. There are, in the literature, a number of exact or asymptotically exact solutions obtained by this method. Here we will merely sketch the basic ideas behind the method as needed for obtaining model-dependent solutions of the relaxation function $R(t)$.

In a space S realized by KSP, $A(t)$ is a vector, whose norm is an invariant of time, i.e. $\|A(t)\| = \|A(0)\|$. Then, as $t \to t'$, $A(t)$ can change only the direction. The change of the direction is determined by the Heisenberg equation of motion; hence, it is model dependent. Because the length is fixed, the trajectory of $A(t)$ is a continuous line drawn on the hypersurface of space S. Note that since S is not a Euclidean space, its shape is not necessarily a hypersphere. The shape will depend explicitly on models themselves. The space is spanned by d basis vectors: $f_0, f_1, \ldots, f_{d-1}$, which satisfy the orthogonality $(f_m, f_{m'}) = 0$ if $m' \neq m$.

These basis vectors may be constructed by the prescriptions of the method of recurrence relations. The Gram–Schmidt process is possible but not helpful to do so as shown elsewhere. The possible values for d are $2, 3, \ldots, \infty$. If $d < \infty$, $A(t)$ is periodic, hence the limit $T \to \infty$ is not possible, i.e. not irreversible. If $d \to \infty$, $A(t)$ is not periodic; hence, $T \to \infty$ is possible. If $N < \infty$, where N is the number of particles in a system, $d < \infty$. If $N \to \infty$, $d < \infty$ or $d \to \infty$. For ergodic behaviour, $d \to \infty$ is a necessary condition since then irreversibility is possible. $N \to \infty$ is also necessary if otherwise irreversibility does not develop.

Let us introduce $r(t) = (A(t), A)/(A, A)$. The normalized relaxation function has a continued fraction representation in the form

$$\tilde{r}(z) = 1/z + \Delta_1/z + \cdots + \Delta_{d-1}/z, \tag{7}$$

where $\Delta_{m+1} = \|f_{m+1}\|/\|f_m\|$, $0 \leqslant m \leqslant d-1$, known as the *recurrants* in the theory. Since these recurrants are made up of the basis vectors, the shape of the realized model-specific space is reflected in $\tilde{r}(z)$. The hypersurface is denoted by $\sigma = (\Delta_1, \Delta_2, \ldots)$.

Now turning to EH, we let $d \to \infty$ on the rhs of (7) and then let $z \to 0$ to obtain an expression for our ergodic condition (5): $\tilde{r}(z = 0) \equiv W$,

$$W = \frac{\Delta_2 \cdot \Delta_4 \cdot \Delta_6 \ldots}{\Delta_1 \cdot \Delta_3 \cdot \Delta_5 \ldots}, \tag{8}$$

an infinite product like the famous one by Wallis. We can now restate our ergodic condition as $0 < W < \infty$, or $\tilde{r}(0)$ or $\int_0^\infty r(t)\,\mathrm{d}t$ is finite. The three different ways are all equivalent, of course, but depending on a problem we may find one way more easily calculated than the others.

We can provide general physical mechanisms when EH is valid and when it is not. If the energy transmitted to a system by a probe in an inelastic scattering process is localized to a finite region of space or to a collective mode, then that system would not be ergodic, signified by $W = 0$. If the energy is taken up by a dominant component of a system, it would also not be ergodic signified by $W = \infty$. If the transmitted energy results in a coherent translation of *all* particles in a system, W will turn out to be finite and the system ergodic.

Before turning to the electron gas models, we shall first illustrate three equivalent ways of determining W. Suppose the hypersurface $\sigma = (1, 4, 9, 16, 25, \ldots)$ in some dimensionless units for simplicity. For this set of recurrants, $r(t) = \operatorname{sech} t$. The three different ways are

$$W = \frac{2^2 \cdot 4^2 \cdot 6^2 \ldots}{1^2 \cdot 3^2 \cdot 5^2 \ldots} \tag{9a}$$

$$\tilde{r}(0) = \int_0^\infty \operatorname{sech} t\,\mathrm{d}t \tag{9b}$$

$$\int_0^\infty r(t)\,\mathrm{d}t = 1/2\{\Psi(z/4 + 3/4) - \Psi(z/4 + 1/4)\}|_{z=0}, \tag{9c}$$

where $\Psi(z) = \mathrm{d}/\mathrm{d}z \log \Gamma(z)$. Each of them yields $\pi/2$, the first being Wallis' infinite product.

4. Density fluctuations in an electron gas

The general model for the electron gas is defined by the Hamiltonian

$$H = \sum_k \varepsilon_k a_k^+ a_k + \frac{1}{2}\sum_{k \neq 0} v_k \rho_{-k} \rho_k, \tag{10}$$

where a_k^+ and a_k are respectively the creation and annihilation operators at wave vector k, v_k is the Coulomb interaction and $\rho_k = \sum_q a_q^+ a_{k-q}$ is the density fluctuation operator. For this model, we shall take the dynamical variable to be the density fluctuation operator $A = \rho_k$. We shall consider $k \ll k_F$ or $k \gg k_F$, k_F is the Fermi wave vector (more simply $k \to 0$ or ∞ measuring in units of k_F), and the space dimensions D to be 1, 2 or 3, for some of which we know *asymptotically exact* forms for the relaxation function. Since our evaluations of W will be for this dynamical variable, we can say whether a system is ergodic only with respect to it, not about a system itself generally.

4.1. 1D ideal gas and Tomonaga–Luttinger model in the ground state

At long wavelengths, a perturbation of the ground state results in a small oscillation about the Fermi surface [23]. Hence $d < \infty$. It is thus not ergodic with respect to this variable.

4.2. 2D ideal gas in the ground state

At long wavelengths and if $v_k = 0$, it is known that $d \to \infty$ [24]. The hypersurface is nearly hyperspherical: $\sigma = (2, 1, 1, 1, \ldots)$, in units of $4k^2\varepsilon_F^2 = 1$, ε_F is the Fermi energy. This gives $r(t) = J_0(2t)$, the Bessel function. We readily obtain

$$W = \frac{1.1.1.1\ldots}{2.1.1.1\ldots} = \frac{1}{2}.$$

The same value is obtained by the integral of $J_0(2t)$ and also from

$$\tilde{r}(z) = 1\Big/\sqrt{z^2+4}.$$

This model is ergodic with respect to the density fluctuation variable. As $k \to 0$, the single particles go into a coherent translation. By the thermodynamic equivalence, we can argue that the same would apply to an ideal Bose gas in 2D [25].

4.3. Nonideal quasi-2D gas at the ground state

If $v_k = 2\pi e^2/k$, at long wavelengths the hypersurface is changed only slightly from the ideal one [24]: $\sigma = (2\lambda, 1, 1, 1, \ldots)$, $\lambda = 1 + (\omega_{pl}/k)^2$, ω_{pl} is the classical plasma frequency. Hence, $W = 1/2\lambda$, also found from

$$\tilde{r}(z) = 1/[(1/\lambda - 1)z + \lambda(z^2+4)^{1/2}].$$

Thus if $\lambda < \infty$, this model is ergodic with respect to $\rho_{k\to 0}$. At the lowest values of k, $\lambda \to \infty$ because of the interaction; thus $W \to 0$. The model ceases to be ergodic as localization sets in. The transmitted energy goes into exciting the plasmon mode, leaving aside the single-particle spectrum.

4.4. 3D ideal gas at the ground state

At the long wavelengths, the hypersurface shows a regular structure of the form $\sigma = (1/3, 4/3.5, 9/5.7, 16/7.9, \ldots)$ in units $4k^2\varepsilon_F^2 = 1$ [26]. The infinite product although seemingly intricate appears to be reducible to Wallis' form. But it can be determined directly from

$$\tilde{r}(z=0) = \arctan 1/z|_{z=0} = \pi/2.$$

That is, $W = \pi/2$. Evidently Wallis' infinite product has another equivalent representation. It is ergodic as the $2d$ ideal gas at long wavelengths.

4.5. 3D nonideal gas at the ground state

If $v_k = 4\pi e^2/k$, the hypersurface at long wavelengths shows a change only in the first recurrant [26]. That is, $1/3 \to \lambda/3$, where now $\lambda = 1 + 3\omega_{\rm pl}^2$, $\omega_{\rm pl}$ is the classical plasma frequency in 3D. Thus, we can readily determine that $W = \pi/2\lambda$. The ergodic behaviour here is similar to what we have shown for the 2D nonideal gas.

4.6. 3D nonideal gas at short wavelengths

If $v_k = 4\pi e^2/k$, the hypersurface when $k \gg k_{\rm F}$ has a regular structure for one special value of r_S at about 3.5 [27], from which we obtain

$$W = \frac{(2.1).(2.2).(2.3)\ldots}{2s.2(s+1).2(s+2)\ldots},$$

where $s = \frac{3k^2}{16}\langle {\rm KE}\rangle|_{rS=3.5} = 0.2568k^2$, k in units of $k_{\rm F}$, $\langle {\rm KE}\rangle$ is the average kinetic energy in units of $\varepsilon_{\rm F}$ the Fermi energy. Since it is more difficult to handle the infinite product, we turn to

$$\tilde{r}(z) = \frac{z}{2\Gamma(s)}\int_0^\infty \frac{{\rm e}^{-u}u^{s-1}\,{\rm d}u}{u + z^2/2}.$$

We observe that $\tilde{r}(0) = 0$ if the integral is finite. It is finite if $s > 1$, approximately $k > 2k_F$. At these very short wavelengths the system ceases to be ergodic, going into the localization limit. It is not difficult to see the underlying physics. At these wave vectors the energy transfers are a result of deep inelastic scattering processes. They probe very small regions of space by confining the probe energy therein. When not delocalized, space averages are no longer related to time averages.

4.7. 2D classical OCP with log potential

Only for a few special systems the hypersurface shows a regular structure with which to determine $r(t)$. For most it would not likely be possible. But one can determine whether d is finite or not. If $d \to \infty$, then one can go to Kubo's condition directly to test ergodicity. Classical OCP gases are possible examples.

One can generally write the frequency ω dependent susceptibility $\chi_k(\omega)$ in terms of the ideal one $\chi_k^o(\omega)$ as follows [28]:

$$\chi_k(\omega) = \chi_k^o(\omega)\big/\big[1 - v_k(1 - G_k(\omega))\chi_k^o(\omega)\big],$$

where $G_k(\omega)$ is the dynamic local field, an undetermined quantity. Now $\chi_k^o(\omega) = V(2\omega/k)\chi_k^o$, where χ_k^o is the ideal static susceptibility and V is the Vlasov function. The Vlasov function behaves as $V(x \to 0) = 1 - 0(x^2)$, where $x \to 0$ means $\omega \to 0$ if k is fixed at some finite value. One does not in general know the ω behaviour in $G_k(\omega)$. But in 2D with a log potential and at $\Gamma = 2$, using the pair correlation function due to Jancovici [29], we know that

$$G_k(\omega) = G_k + 0(\omega^2).$$

Thus, if $k \neq 0$ or ∞, $\chi_k(\omega = 0) = \chi_k$, hence ergodic.

5. Concluding remarks

We have demonstrated through a number of physical examples that it is possible to perform time averages with which to compare the ensemble averages. It was an approach first seriously undertaken by Kubo. But his work was limited owing to the fact that at his time a general solution to the Heisenberg equation was unavailable. We have shown that Kubo's ergodic condition is necessary but not sufficient. Irreversibility and even the thermodynamic limit themselves alone cannot determine whether a system is ergodic. The sufficient condition for ergodicity is finiteness of a uniquely defined infinite product denoted by W. The models of the electron gas were used to illustrate this particular property.

Two important remarks should be made. As perhaps evident from our examples in section 4, a model is not absolutely ergodic or not ergodic. It is with respect to a dynamical variable with which time averages are being made. While at long wavelengths a dynamical variable may lead to an ergodic behaviour, at short wavelengths it may not. Although not shown in this work, one dynamical variable may make a model ergodic while another may make the same model not so. Our examples are limited to a few *exact or asymptotically exact* situations. But it is clear that if a model is ergodic with respect to a variable say at long wavelengths, it is likely so also at somewhat different wavelengths. If, similarly, a model is proved to be ergodic at the ground state, it is likely to remain so at near the ground state.

The ergodic hypothesis (EH), as stated at the outset, is a deep subject. Although our analysis pertains only to the response functions, we believe we have made a first step in providing both mathematical means and physical understanding with which to study EH in many-body systems.

EH is also found in, e.g., nonlinear dynamics and chaos [30–33]. Here the systems have few degrees of freedom (often one or two) and they are not in thermal equilibrium. Also the dynamics are driven usually by ersatz equations of motion, some (discretized) self-similar maps. The proper equation of motion for a many-body system, the Heisenberg equation, is not self-similar except for some trivial models. Given these vast differences, it does not seem likely that the ergodicity in nonlinear dynamics, say, could have much to do with the ergodicity in statistical thermodynamics, the subject of our study. To our knowledge there are no proofs that they are even equivalent.

Acknowledgments

A portion of this work was accomplished at Korea Institute for Advanced Study, Seoul. The author thanks Dr M W Kim, director of the Institute for his hospitality.

References

[1] Landsberg P T 1978 *Thermodynamics and Statistical Mechanics* (New York: Dover) p 139
Pathria R *Statistical Mechanics* (Oxford: Pergamon) p 38
[2] Faquhar I E 1964 *Ergodic Theory in Statistical Mechanics* (New York: Wiley)
[3] Dorfman J R 1999 *An Introduction to Chaos in Nonequilibrium Statistical Mechanics* (New York: Cambridge University Press) p 7
[4] ter Haar D 1960 *Element of Statistical Mechanics* (New York: Holt) p 333
[5] Khintchin A I 1949 *Mathematical Foundations of Statistical Mechanics* (New York: Dover)
Kubo R 1957 *J. Phys. Soc. Japan* **12** 570
[6] Lee M H 1982 *Phys. Rev.* B **26** 2547
Lee M H 1982 *Phys. Rev. Lett.* **49** 1072
[7] Lee M H 2001 *Phys. Rev. Lett.* **87** 250601
[8] Golden K J and Kalman G J 2000 *Phys. Plasmas* **7** 14
[9] Macchi A, Maradudin A A and Tognetti V 1996 *Phys. Rev.* B **53** 5363
[10] Röpke G, Selchow A, Wierling A and Reinholz H 1999 *Phys. Lett.* A **260** 365
[11] Allegrini P, Grigolini P and Rocco A 1997 *Phys. Lett.* A **233** 309
[12] Aksenov V L, Bobeth M, Plakida N M and Schreiber J 1087 *J. Phys. C: Solid State Phys.* **20** 375
[13] Kuzemsky A L 2005 *Int. J. Mod. Phys.* B **19** 1029

[14] Norman G E 2002 *AIP Conf. Proc.* **643** 175, 442
[15] Belyi V V 2004 *Phys. Rev.* E **69** 017104
[16] Yulmetyev R M, Mokshin A V and Hänggi P 2003 *Phys. Rev.* E **68** 051201
Mokshin A V, Yulmetyev R M and Hänggi P 2005 *Phys. Rev. Lett.* **95** 200601
[17] Morgando R, Oliviera F A, Bartrouni G G and Hansen A 2002 *Phys. Rev. Lett.* **89** 100601
[18] Daligault J and Murillo M S 2003 *Phys. Rev.* E **68** 015401
[19] Prosen T 1998 *Phys. Rev. Lett.* **80** 1808
[20] Silva Nunez M E and Florencio J 2003 *Phys. Rev.* B **68** 014406
[21] Sawada I 2002 *Physica* A **315** 14
[22] Srokowsky T 2000 *Phys. Rev. Lett.* **85** 2232
[23] Lee M H, Hong J and Sharma N L 1984 *Phys. Rev.* A **29** 1561
[24] Lee M H and Hong J 1982 *Phys. Rev. Lett.* **48** 634
[25] Lee M H 1997 *Phys. Rev.* E **55** 1518
[26] Lee M H and Hong J 1984 *Phys. Rev.* B **30** 6756
[27] Hong H and Lee M H 1993 *Phys. Rev. Lett.* **70** 1972
[28] Hong J, Park J and Lee M H 1989 *Phys. Rev.* B **40** 1528
[29] Jancovici B 1981 *Phys. Rev. Lett.* **46** 389
[30] Schuster H G 1984 *Deterministic Chaos* (Weinheim: Physik Verlag)
[31] Sinai Ya G and Vul E B 1981 *Physica* D **2** 3
[32] Buminovich L A and Sinai Ya G 1980 *Commun. Math. Phys.* **78** 247
[33] Zaslavsky G M 1981 *Phys. Rep.* **80** 157
Zaslavsky G M 1978 *Phys. Lett.* A **69** 145