Why Irreversibility Is Not a Sufficient Condition for Ergodicity

M. Howard Lee*

Department of Physics and Astronomy, University of Georgia, Athens, Georgia 30602, USA
Korea Institute for Advanced Study, Seoul 130-012, Korea

(Received 20 November 2006; revised manuscript received 16 January 2007; published 8 May 2007)

Khinchin's theorem of ergodicity is examined by means of linear response theory. The resulting ergodic condition shows that, contrary to the theorem, irreversibility is not a sufficient condition for ergodicity. By the recurrence relations method, we prove that irreversibility is broader in scope than ergodicity, showing why it can only be a necessary condition for ergodicity.

DOI: 10.1103/PhysRevLett.98.190601 PACS numbers: 05.70.Ln, 05.20.−y, 05.30.−d, 05.40.−a

1. Introduction.—Khinchin is an important name in ergodic theory, a branch of mathematics, tracing its origin to Boltzmann's ergodic hypothesis. Unlike most others in this field, Khinchin formulated his work in terms of correlation functions, the language of statistical mechanics. It is thus no surprise to find that his work has made a deep impact in physics, e.g., Kubo [1]. Khinchin's theorem states that a variable A is ergodic if the autocorrelation function of A is "irreversible." We contend that this theorem cannot be completely correct if A refers to a Hermitian system and if the averages of A are given by the linear response theory of inelastic scattering processes. We prove that the irreversibility of A is only a necessary condition, not a sufficient one, for the ergodicity of A in such a system.

2. Khinchin's theorem [2].—The theorem states that, for a classical system, the ergodicity of a dynamical variable A in thermal equilibrium

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\langle A(t)\rangle dt = \langle A\rangle \tag{1}$$

must hold if the autocorrelation function of A satisfies the relation

$$\lim_{t\to\infty}\langle A(t)A(0)\rangle = \langle A\rangle\langle A\rangle, \tag{2}$$

where $A = A(0)$ and the angular brackets denote an ensemble average, further defined in section 3. Thus, according to Khinchin, (2) implies (1). Kubo asserts after Khinchin that the theorem may be reversed. That is true because, as we shall show, (1) requires (2). The two are, however, not one to one as Kubo seemed to have believed.

3. Linear response theory and time averages.—We first consider a quantum system and thereafter a classical one. Let a system denoted by H be Hermitian. We turn on a weak field $h(t)$ at a remote past $t = -T$, $T \to \infty$. The total energy at t is given by

$$H'(t) = H(A) + h(t)A. \tag{3}$$

According to linear response theory [1],

$$\langle A(t)\rangle_{H'(t)} = \langle A(t)\rangle_H + \int_{-T}^{t} h(t')\chi_A(t-t')dt', \tag{4}$$

where $\langle\ldots\rangle_{H'}$ and $\langle\ldots\rangle_H$ mean ensemble averages with the density matrices of H' and H, respectively, and $\chi_A(t-t')$ is a response function, depending on H, not H', assumed to be both causal and stationary. Let us take a time average (TA) on (4): noting that for a Hermitian system, $\langle A(t)\rangle_H = \langle A\rangle_H$ [3], we obtain

$$I_{\rm TA} \equiv \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\langle A(t)\rangle_{H'(t)}dt$$
$$= \langle A\rangle_H + \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{t} h(t')\chi_A(t-t')dt'dt. \tag{5}$$

Suppose $h(t) = h$, a constant field. Then the right-hand side of (5) becomes

$$I_{\rm TA} = \langle A\rangle_H + h\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{t}\chi_A(t-t')dt'dt. \tag{6}$$

Let us assume (after Khinchin) the ergodicity of A:

$$I_{\rm TA} = \langle A\rangle_{H'}, \tag{7}$$

where $H' = H + hA$. By linear response theory, $\langle A\rangle_{H'} = \langle A\rangle_H + h\chi_A$, where χ_A is the static response function [1]. Thus assuming the ergodicity of A is the same as assuming:

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{t}\chi_A(t-t')dt'dt = \chi_A. \tag{8}$$

Thus if (8) is proved, (1) is proved to order h, which is sufficient. Now we already have proved that (8) is valid if the following condition holds [4]: $0 < W < \infty$, where

$$W = \int_0^\infty R_A(t)\,dt, \tag{9}$$

where $R_A(t)$ is the autocorrelation function of A

$$R_A(t) = (A(t), A), \tag{10}$$

where $R_A(0) = \chi_A$. The inner product is defined as [5]: if A

and B are Hermitian operators,

$$(A, B) = 1/\beta \int_0^\beta \langle e^{\lambda H} A e^{-\lambda H} B \rangle d\lambda - \langle A \rangle \langle B \rangle, \quad (11)$$

where β is the inverse temperature and the brackets are ensemble averages with the density matrix of H. In the classical limit ($\beta \to 0$), the first term on the right-hand side of (11) becomes $\langle AB \rangle$ [6]. If A is ergodic, i.e., W finite, (9) requires $R_A(t = \infty) = 0$, which we call the *irreversibility* of A. It is possible that $W = 0$ or ∞ while $R_A(t = \infty) = 0$ [9]. Thus irreversibility alone does not ensure ergodicity. That is why irreversibility is a necessary but not sufficient condition for ergodicity. If A ceases to be irreversible (for example, as the critical region is approached), it ceases to be ergodic. It underlines the importance of determining irreversibility from a general consideration.

4. Irreversibility.—We shall now prove that the irreversibility of A in a Hermitian model is a general property of A. The proof is by the recurrence relations method, now well established [8,10,11]. We shall state it without proof. Let S denote an inner product space of self-adjoint operators A and B, having the properties among others: $(A, B) = (B, A)$ and $(A, OB) = (O^* A, B)$, where O is an operator in S and $*$ means Hermitian conjugation. Let $A(t)$ be a vector in S, where $A(t) = e^{iHt} A e^{-iHt}$, $A = A(t = 0)$, and $\hbar = 1$. If H is Hermitian, the norm of $A(t)$ in S is an invariant of the time t:

$$(A(t), A(t)) = (A, A). \quad (12)$$

It means that as t evolves, $A(t)$ can change only its directions in S. It may be expressed by an orthogonal expansion,

$$A(t) = \sum_{\nu=0}^{d-1} a_\nu(t) f_\nu, \quad (13)$$

where $\{f_\nu\}$ is a complete set of basis vectors that span S of d dimensions, i.e., $(f_\nu, f_{\nu'}) = 0$ if $\nu' \neq \nu$, and $a_\nu(t)$ is the magnitude of the projection of $A(t)$ onto f_ν at t, hence a real function of t. Here d may be finite or infinite. Let t evolve from $t = 0$. Then it is natural to choose $f_0 = A(t = 0) = A$. The others we construct by a recurrence relation, see below. With this choice, $a_0(t) = (A(t), A)/(A, A) = R_A(t)/R_A(0) \equiv r_A(t)$. Thus, $a_\nu(0) = 1$ if $\nu = 0$ and $= 0$ if otherwise (boundary condition). By Schwarz inequality, $|a_0(t)| \leq 1$. By (12) and (13), we obtain the Bessel equality (BE):

$$\sum_{\nu=0}^{d-1} [a'_\nu(t)]^2 = 1, \quad (14)$$

where ${a'}_0 = a_0$ and $a'_\nu = \sqrt{\Delta_1 \Delta_2 \ldots \Delta_\nu} a_\nu$, where $\Delta_\nu = (f_\nu, f_\nu)/(f_{\nu-1}, f_{\nu-1})$, $\nu = 1.2, \ldots d - 1$. By Schwarz inequality, $|{a'}_\nu(t)| \leq 1$, also implied by BE.

Now we *realize* S by (11) and call it S' [12]. It is still an inner product space, but is H specific meaning H dependent. In S', $\{f_\nu\}$ can be constructed starting with $f_0 = A$ by a recurrence relation known as RR1, which is H specific [13]. Thus, Δ_ν's are also H specific. RR1 implies a recurrence relation for $\{a_\nu\}$, known as RR2 [8]:

$$\Delta_1 a_1 = -d/dt a_0 \quad (15a)$$

$$\Delta_{\nu+1} a_{\nu+1} = -d/dt\, a_\nu + a_{\nu-1}, \qquad \nu = 1, 2, \ldots d - 1. \quad (15b)$$

RR2 (15a) and (15b) itself is realized if the values of Δ_ν's are specified. The solutions for $\{a_\nu\}$ are per force H specific and thus unique. They have the following important properties: (a) linearly independent if $0 < t < \infty$; (b) bounded and analytic everywhere in real t. Solutions for $\{a_\nu\}$ which meet these properties are termed *admissibles*. An admissible satisfies BE collectively and Schwarz inequality individually.

If $d < \infty$, $a_0(t + lt_1) = a_0(t)$, $l = 1, 2, \ldots$, for some t_1 [14]. Thus $a_0(t = \infty) \neq 0$, not irreversible. Henceforth we consider $d \to \infty$ only, replacing index ν by n. If $t \to 0$, $a_0 \to 1$ (see the boundary condition). As t increases, the amplitude may decrease. Let $t \to \infty$. We assume (Assumption A) that for all admissibles, $d/dt\, a_0(t = \infty) = 0$. Assumption A is justified by property b: there are no singularities in a_0 as $t \to \infty$. Since $|a_0(t = \infty)| \leq 1$, it must reach a limiting value asymptotically with zero slope.

Then, by Assumption A on (15a), $a_1(t = \infty) = 0$. By (15b) with $\nu = n = 1$, $\Delta_2 a_2 = -d/dt a_1 + a_0 = a_0$ since, if $a_1(\infty) = 0$, $d/dt a_1(\infty) = 0$. If $\nu = n = 2$, $\Delta_3 a_3 = -d/dt\ a_2 + a_1 = 0$ since $d/dt a_2 = d/dt a_0/\Delta_2 = 0$. A continuation shows that all odd ones are zero and all even ones are connected: as $t \to \infty$, for $n = 1, 2, \ldots$,

$$a_{2n-1} = 0, \quad (16a)$$

$$\Delta_2 \Delta_4 \ldots \Delta_{2n} a_{2n} = a_0. \quad (16b)$$

When (16a) and (16b) are substituted in BE (14), we obtain (now writing r for a_0): $|r(\infty)| = 1/\sqrt{K}$, where

$$\begin{aligned} K &= 1 + \Delta_1/\Delta_2 + \Delta_1\Delta_3/\Delta_2\Delta_4 + \ldots \\ &\quad + (\Delta_1\Delta_3 \ldots \Delta_{2n-1})/(\Delta_2\Delta_4 \ldots \Delta_{2n}) + \ldots \\ &\equiv 1 + \sum_{n=1}^{\infty} k_{2n}. \end{aligned} \quad (17)$$

If the K series converges to a finite number K_0 say, $|r(\infty)| = 1/\sqrt{K_0} > 0$, hence not irreversible. If it diverges, $r(\infty) = 0$ and irreversible. We showed that $W = 1/k_{2n}$, $n \to \infty$ [4]. The very "last" term in the K series is thus $1/W$ [15]. If A is ergodic, i.e., its W is finite, its K series (whose k_∞ is finite) must diverge. It is necessarily irreversible. If nonergodic by $W = 0$, the K series (whose $k_\infty = \infty$) surely diverges and is irreversible. If nonergodic by $W = \infty$, its K series (whose $k_\infty = 0$) may or may not diverge. This is all consistent with the ergodic condition. We now show that $|r(\infty)|$ can be determined independently.

First, consider A in a system of one interaction constant, e.g., J (exchange constant), termed *homogeneous*. For simplicity, we look at a delocalization process [16] on a lattice, rather than in a fluid, although the outcome is basically similar, in some cases even isomorphic [7(a)]. Let it be a linear chain of spins coupled between nn's by an exchange constant J (e.g., XY model), where the spin sites are labeled $0, 1, 2, \ldots N$, $N \to \infty$. If the spin at site 0 is perturbed by an external means, the perturbation energy will spread down the chain, a process termed the delocalization of the perturbation energy. Roughly put, Δ_n measures the delocalization of the perturbation energy from site $n-1$ to site n, having its process started at site 0 [7(b)]. It may be given as: $\Delta_n = J^2 g(n)$, where $g(n)$ has H specific details. A known $g(n)$ means a specific shape of S' and a corresponding admissible [10,11].

The site-to-site delocalization does not end since $d \to \infty$. (If $d < \infty$, $\Delta_d = 0$, resulting in a periodic admissible [14].) For the purpose of (17), we actually do not have to know the precise form of $g(n)$. We need to know only the ratio or rate Δ_n/Δ_{n-1} as $n \to \infty$, with which to determine whether the K series diverges, not how it diverges. As $n \to \infty$ (i.e., sufficiently far removed from the site of the perturbation), the rate in a homogeneous system should become n independent. Since the delocalization of the perturbation energy is continuous in a linear homogeneous system, the rate should approach a constant. That is, $[\Delta_n/\Delta_{n-1} - \Delta_{n-1}/\Delta_{n-2}]_{n\to\infty} = 0$. Thus we assume (Assumption B) that if site n is sufficiently far removed from site 0,

$$\Delta_n/\Delta_{n-1}|_{n\to\infty} = 1 + 0(1/n). \qquad (18)$$

To justify it fully we would need to compare f_{n-1} versus f_n for a given H [7(b)]. There is now a fairly large body of exactly known $g(n)$'s for a variety of homogeneous lattice and fluid models, both quantum and classical [7,10,11]. We list a few elementary ones: $g(n) = 1$ (hypersphere), $g(n) = n$ or n^2 (hyperellipsoid), $g(n) = 4n^2/(4n^2 - 1)$ (asymptotically hypersphere), etc. They all satisfy Assumption B. For a homogeneous system, by Assumption B: $K \sim d/2$, $d \to \infty$, where d is the dimensions of S'. Thus the K series diverges (as already forecast by k_∞) and, by (17), $r(\infty) = 0$. It is irreversible. For all Hermitian *homogeneous* systems belonging to S' of $d = \infty$, we conclude by Assumption B that irreversibility is a general property of A [17–19]. All known admissibles show this property without exception, e.g., secht, $\exp -t^2/2$, $J_0(2t)$, $J_1(2t)/t$, $j_0(2t)$ (sph. Bessel), $M(s, 1/2, -t^2/4)$ (Kummer) [7,10,11].

Assumption B can break down if $\Delta_1 \to 0$ (others finite) in some domains (critical region, Brownian limit). Then, $k_{2n} \to 0$ for every n. See (17). Thus $K \to K_0 = 1$ and $|r(\infty)| = 1$. When these domains are entered, irreversibility is lost [4]. If irreversibility is lost, ergodicity is also lost. See (9). Thus any analysis of critical behavior or Brownian motion [20] which assumes ergodicity raises questions.

An *inhomogeneous* system has two or more interaction constants, e.g., Jx and Jy (anisotropic XY model), J and B (transverse Ising model). For a two-constant system we have a new parameter γ say, e.g., $\gamma = (Jy/Jx)^2$ or $(B/J)^2$, which thus defines two regions: I ($\gamma < 1$) and II ($\gamma > 1$), separated by $\gamma = 1$, a homogeneous point. Assumption B would not apply here. Instead of another, we consider a simple structure, realized exactly in a two-constant model: $\Delta_{2n-1} = \gamma$ and $\Delta_{2n} = 1$, $n = 1, 2 \ldots$, and $0 < \gamma < \infty$. By (17)

$$K = 1 + \gamma + \gamma^2 + \ldots + \gamma^n + \ldots. \qquad (19)$$

If $\gamma < 1$ (I), K converges to $K_0 = 1/(1-\gamma)$ and $|r(\infty)| = \sqrt{(1-\gamma)} > 0$. If $\gamma > 1$ (II), K diverges and $r(\infty) = 0$. By $W = 1/k_{2n}$, $n \to \infty$, where $k_{2n} = \gamma^n$, $W = \infty$ in I and $W = 0$ in II. In both regions it is not ergodic. If $\gamma = 1$ (homogeneous pt.), it gives a hyperspherical S', for which we already know $r(t) = J_1(2t)/t$ [21]. We see that $r(\infty) = 0$ (irreversible) and $W = \int_0^\infty J_1(2t)dt/t = 1$ (ergodic). They are precisely borne out by the K series: when $\gamma = 1$, $K = \infty$ (Assumption B satisfied) and $W = 1/k_\infty = 1$.

5. Concluding remarks.—Ergodicity is represented by just one term in the series whereas irreversibility by the entire series [15]. Thus the two are not one to one. This nonisomorphic relationship is the reason why irreversibility can only be a necessary condition for ergodicity.

In a homogeneous system, irreversibility is a rule except in anomalous domains where Assumption B breaks down. In an inhomogeneous system the irreversibility exists if $\gamma > 1$ and does not if $\gamma < 1$. Now the value of γ needs not always come fixed. It could be varied by adjusting the strengths of the interaction constants (e.g., by changing B while fixing J in the transverse Ising model). Thus by exercising this extra degree of freedom, one could create or destroy the irreversibility. The system itself, however, remains nonergodic whether the irreversibility exists or not.

I wish to thank Professors H. Park and M. W. Kim of KIAS, Seoul for their kind hospitality where a portion of this work was accomplished.

*Email address: MHLee@uga.edu

[1] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
[2] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949); also see p. 577 of Ref. [1].

[3] In an ordered state $\langle A \rangle_H \neq 0$. But this term does not contribute to the ergodicity condition on W (9).

[4] M. H. Lee, Phys. Rev. Lett. **87**, 250601 (2001); Physica (Amsterdam) **314A**, 583 (2002); **365A**, 150 (2006); J. Phys. A **39**, 4651 (2006).
[5] M. H. Lee, Phys. Rev. Lett. **49**, 1072 (1982).
[6] If $\beta \to 0$ (classical limit), $R(t \to \infty) = 0$ thus reduces to (2). Classical models and quantum models in the classical limit would take the classical form of the inner product: $R(t) = \langle A(t)A\rangle - \langle A\rangle\langle A\rangle$. See [7(b)]. The recurrence relations method [8] yields $R(t)$ for a classical model by replacing commutators with Poisson brackets (pb), e.g., $d/dtA(t) = i[H, A] \to -[H, A]_{\text{pb}}$. There now exist in the literature a number of classical solutions thus obtained. See, e.g., [7(a)]. Thus our formalism given here for a quantum system applies to both the classical limit of a quantum system and a classical system.
[7] (a) M. H. Lee, J. Florencio, and J. Hong, J. Phys. A **22**, L331 (1989); (b) J. Florencio, Jr. and M. H. Lee, Phys. Rev. B **35**, 1835 (1987), see Eqs. (3.2a–c), (3.9a–d).
[8] M. H. Lee, Phys. Rev. B **26**, 2547 (1982); J. Math Phys. **24**, 2512 (1983).
[9] If $R_A(t) = J_0(t) - J_4(t)$, $W = 0$. If $R_A(t) = [J_0(t)]^2$, $W = \infty$. Both autocorrelation functions are evidently irreversible. They are actually realizable in many body models [4].
[10] For reviews, see A. S. T. Pires, Helv. Phys. Acta **61**, 988 (1988); V. S. Viswanath and G. Müller, *Recursion Method* (Springer-Verlag, Berlin, 1994); U. Balucani, M. H. Lee, and V. Tognetti, Phys. Rep. **373**, 409 (2003); A. V. Mokshin and R. M. Yumetyev, *Microscopic Dynamics of Simple Liquids* (Center for Innovative Technology, Kazan, 2006) (in Russian).
[11] For applications by others, see, e.g., C. Lee and S. I. Kobayashi, Phys. Rev. Lett. **62**, 1061 (1989); D. Vitali and P. Grigolini, Phys. Rev. A **39**, 1486 (1989); S. G. Jo, K. H. Lee, S. C. Kim, and S. D. Choi, Phys. Rev. E **55**, 3676 (1997); I. Sawada, Phys. Rev. Lett. **83**, 1668 (1999); J. Dauligault and M. S. Murillo, Phys. Rev. E **68**, 015401 (2003); V. V. Belyi, Phys. Rev. E **69**, 017104 (2004); A. V. Mokshin, R. M. Yulmetyev, and P. Hänggi, Phys. Rev. Lett. **95**, 200601 (2005); Z.-Q. Liu, X.-M. Kong, and X.-S. Chen, Phys. Rev. B **73**, 224412 (2006); A. V. Mokshin, R. M. Yulmetyev, R. M. Khusnutdinoff, and P. Hänggi, J. Phys. Condens. Matter **19**, 046209 (2007); A. Wierling, H. Reinholz, G. Röpke, and J. Adams, Contrib. Plasma Phys. **45**, 441 (2005).
[12] An abstract inner product space such as S is a general space. It can be made specific if a specific meaning is assigned to an inner product, a process known as realizing an abstract space. See [5].
[13] S' is an inner product space realized by (11), sometimes called Kubo's scalar product. The orthogonalization process for this realized space is given by a recurrence relation (known as RR1): $f_{\nu+1} = i[H, f_\nu] + \Delta_\nu f_{\nu-1}$, $\nu = 0, 1, 2 \ldots d-1$, with $f_{-1} \equiv 0$ and $\Delta_1 \equiv 1$. See [5]. Given f_0, all higher ones can be obtained one by one by RR1. Observe that H explicitly appears in RR1; hence, all f_ν's are H specific except f_0.
[14] If $d = 1$, (13) and (15a) give: $A(t) = a_0(t)f_0$ and $d/dta_o(t) = 0$, respectively. Hence, $A(t) = A$, a constant of motion. If $d = 2$, (15a) and (15b) give: $\Delta_1 a_1 = -d/dta_0$ and $-d/dta_1 + a_0 = 0$, respectively. They imply: $a_0 = \cos\sqrt{\Delta_1}t$ and $a_1 = \sin\sqrt{\Delta_1}t/\sqrt{\Delta_1}$, which satisfy BE. If d is finite, a_0 consists of the circular functions only since they come from isolated simple poles in the complex frequency plane. See M. H. Lee, Can. J. Phys. **61**, 428 (1983). For most dynamic models, $d \to \infty$ if $N \to \infty$, where N is the number of particles in H. When $d \to \infty$, the isolated poles can coalesce making $a_0(t)$ nonperiodic (e.g., Bessel functions).
[15] If K is a set, the individual k's are like elements of the set and the irreversibility is a property of this set K. The ergodicity is denoted by one element of the set. Thus one might say that the ergodicity is contained in or belongs to the irreversibility.
[16] In [4] we show that one can understand ergodicity from a perturbation picture: suppose an interacting system is perturbed by an external probe through A. If the perturbation energy becomes delocalized to all coordinates of the system, A is ergodic. If the delocalization process does not reach everywhere in the system, A is not ergodic.
[17] T. Prosen, J. Phys. A (to be published).
[18] M. H. Vainstein, I. V. L. Costa, and F. A. Oliveira, Lect. Notes Phys. **688**, 159 (2006); I. V. L. Costa, R. Morgado, M. V. B. Lima, and F. A. Oliveira, Europhys. Lett. **63**, 173 (2003); L. C. Lapas, I. V. L. Costa, M. H. Vainstein, and F. A. Oliveira, Europhys. Lett. **77**, 37 004 (2007).
[19] P. Grigolini, *Quantum Mechanical Irreversibility and Measurement* (World Scientific, Singapore, 1993), p. 129.
[20] J.-D. Bao, Y.-Z. Zhuo, F. A. Oliveira, and P. Hänggi, Phys. Rev. E **74**, 061111 (2006).
[21] S. Sen, Phys. Rev. B **44**, 7444 (1991); M. H. Lee and J. Hong, Phys. Rev. B **32**, 7734 (1985).